

Doble Grado en Estadística y Economía

Título: Las criptomonedas en el sistema financiero internacional. Un análisis predictivo mediante técnicas de Machine Learning (LSTM)

Autor: Arnau Castillo Mur

Director: Salvador Torra Porras

Departamento: Departamento de Econometría (UB)

Convocatoria: Junio 2020



UNIVERSITAT DE
BARCELONA



UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Facultat de Matemàtiques i Estadística

Resumen y palabras clave

Este trabajo está compuesto de dos partes. En la primera se lleva a cabo una presentación de lo que son las criptomonedas y cuáles son sus principales características. Se pone especial foco en Libra, moneda digital corporativa propuesta por Facebook basada en el uso de la tecnología Blockchain que quiere convertirse en una alternativa más segura y escalable respecto a las criptomonedas y que busca tener un gran impacto en los países menos desarrollados. A continuación, se analizan sus principales características, su funcionamiento, objetivos, competencia y el motivo de su fracaso hasta ahora. Además, se hace un análisis estadístico de las tarifas aplicadas a las remesas de 2011 a 2019 y de los niveles de acceso a servicios financieros de los años 2011, 2014 y 2017 para diferentes regiones del mundo. Esto se hace con el objetivo de analizar las variables que están detrás de la aparición de alternativas a los medios convencionales de transferencia de dinero, como por ejemplo Libra. Con esta información se podrá concluir como de necesarias son estas alternativas y si se está produciendo una convergencia de los países en desarrollo respecto a los desarrollados en las mencionadas variables. Por otro lado, se realiza un estudio de cómo podría ser un sistema financiero basado en criptomonedas, analizando la estabilidad de precios, la eficiencia de pagos, el riesgo sistémico endógeno y la posibilidad de que bancos centrales emitan monedas digitales con una funcionalidad parecida a la de Libra.

En la segunda parte del trabajo se realiza en primer lugar un análisis estadístico de la rentabilidad de una serie de criptomonedas para así poder concluir cual es el riesgo asociado a cada una de ellas. Una vez hecho esto, se hace una distinción entre los modelos estadísticos y los modelos de Machine Learning y se demuestra estadísticamente por qué un modelo perfectamente especificado no tiene por qué ser el que realiza las mejores predicciones, aspecto que se tendrá en cuenta más adelante. En segundo lugar, se analiza la capacidad predictiva del precio de cierre de modelos de Machine Learning como son las redes LSTM y de modelos estadísticos como el modelo ARIMA. Con las predicciones realizadas por estos modelos se analizarán los beneficios que se pueden obtener con una estrategia de trading muy sencilla. Por último, se justifica que motivos hacen a la red LSTM una de las mejores candidatas para la predicción del precio de cierre por encima de otras alternativas como son las redes neuronales recurrentes (RNN).

Palabras clave: Libra, criptomoneda, remesas, RNN, modelo predictivo de criptomonedas, Blockchain, ARIMA, LSTM.

Abstract

This work is composed of two parts. The first one starts with a presentation of what cryptocurrencies are and what are their main characteristics. Special focus is placed on Libra, a Corporate Digital Currency proposed by Facebook based on the use of Blockchain technology that wants to become a more secure and scalable alternative to cryptocurrencies and that seeks to have a great impact in less developed countries. The main characteristics, its operation, objectives, competence and the reason for its failure so far are analyzed. In addition, a statistical analysis is made of the rates applied to remittances from 2011 to 2019 and of the levels of access to financial services for the years 2011, 2014 and 2017 for different regions of the world. This is done with the aim of analyzing the variables that are behind the emergence of alternatives to the means of money transfer, such as Libra. With this information, it will be possible to conclude how necessary these alternatives are and whether there is a convergence of developing countries with respect to developed countries in the aforementioned variables. On the other hand, a study is made of what a cryptocurrency-based financial system could be like, analyzing price stability, payment efficiency, endogenous systemic risk and the possibility that central banks issue digital currencies with functionality similar to that of Libra.

In the second part of the work, a statistical analysis of the profitability of a series of cryptocurrencies is carried out in order to be able to conclude what is the risk associated with each of them. Once this is done, a distinction is made between statistical models and Machine Learning models and it is statistically demonstrated why a perfectly specified model does not have to be the one that makes the best predictions, an aspect that will be considered later. Secondly, the predictive capability of the closing price of Machine Learning models such as LSTM networks and statistical models such as the ARIMA model are analyzed. With the predictions made by these models there is an analysis of the benefits that can be obtained with a very simple trading strategy. Finally, the reason that make the LSTM network one of the best candidates for predicting the closing price above other alternatives such as recurrent neural networks (RNN) is justified.

Key words: Libra, cryptocurrency, remittances, RNN, predictive model of cryptocurrencies, Blockchain, ARIMA, LSTM.

índice

1	Introducción	3
2	Metodología	5
3	Tipos de dinero	8
3.1	Dinero fiat, vales, E-money	8
3.2	Criptomonedas	9
3.2.1	La tecnología blockchain	9
3.2.2	Tipos de Blockchain	15
3.2.3	Diferencia entre criptomonedas, altcoins, stablecoins, tokens, monedas virtuales y monedas digitales	16
3.3	Monedas digitales corporativas: Libra, concepción original	18
3.3.1	La Reserva de Libra	19
3.3.1.1	Funcionamiento y estructura	19
3.3.1.2	Interacción entre las entidades y la reserva	20
3.3.1.3	El cambio de la reserva a lo largo del tiempo	21
3.3.1.4	Política monetaria	21
3.4	Monedas digitales de bancos centrales	22
4	Coexistencia del dinero fiat y las criptomonedas	25
5	¿En qué sistema monetario estaría basado un criptosistema?	27
6	La estabilidad de precios	29
7	La estabilidad financiera de un criptosistema	31
8	Las transferencias de dinero	33
9	¿Por qué Libra no va a ser lo que pretendía?	34
10	M-Pesa, un caso de éxito	37
11	Análisis estadístico de las remesas y su evolución	40
12	Análisis de la entrada y salida de remesas	42
13	Análisis coste (%) de enviar 200 dólares	46
14	Análisis base de datos Findex	50
15	Análisis estadístico de las criptomonedas	57
15.1	Introducción a las criptomonedas usadas	57
15.2	Resumen estadístico criptomonedas	59
15.3	Distribuciones de probabilidad y criterios de selección	60
15.3.1	Distribuciones seleccionadas y parámetros	64

15.3.2	Criterios de información	65
15.3.3	Q-Q Plots	68
15.3.4	P-P Plots	70
15.3.5	Test Bondad de Ajuste	71
15.3.6	VaR plots	71
15.3.7	CVaR plots	72
16	Algoritmos de ML vs modelos estadísticos	75
17	Introducción a la predicción del precio de las criptomonedas	78
17.1	Preparación antes del ajuste de los modelos	79
17.1.1	Redes LSTM	79
17.1.2	Definición de las particiones	81
17.1.3	Modelo ARIMA	82
17.1.3.1	Determinación del orden de integración de las criptomonedas . .	83
17.1.3.2	Aplicación de la diferenciación	85
18	Ajuste del modelo ARIMA	85
19	Las Redes neuronales recurrentes vanilla y LSTM	88
19.1	RNR vanilla	88
19.1.1	Método del gradiente descendiente	91
19.1.2	Formulación de la fase de retropropagación a través del tiempo para caso concreto	93
19.1.3	Formulación de la fase de retropropagación a través del tiempo generalizada	97
19.2	Red LSTM	98
19.2.1	Estructura de la celda LSTM	98
19.2.2	Por qué las LSTM solucionan el problema del gradiente desvaneciente . .	100
19.2.3	Ajuste de la red LSTM	101
20	Ajuste de todos los modelos en el training set	105
20.1	Predicción evolución del precio de cierre	111
21	Conclusiones	117
22	Código	126
22.1	Análisis estadístico de las criptomonedas	127
22.2	Modelo ARIMA y redes LSTM	131
23	Bibliografía	136
23.1	Fuente de los datos	136
23.2	Fuentes Multimedia	137
23.3	Fuentes de artículos académicos y libros	138

1 Introducción

En 2009 se creó la primera criptomoneda, denominada Bitcoin, y desde entonces ha estado en el centro de atención a nivel mundial. Respecto a los métodos tradicionales de pago, el Bitcoin tiene unas características distintivas como son la anonimidad, la alta liquidez, su uso como cobertura contra la inflación en países como Venezuela, la posibilidad de realizar micropagos o los bajos costes asociados a la realización de las transacciones. Está basada en la tecnología blockchain, que consiste en una estructura de datos en la que la información de todas las transacciones realizadas es registrada en una secuencia de bloques. Informalmente se puede entender como un archivo Excel, que está totalmente descentralizado y al que por lo tanto todo el mundo tiene acceso, en el que se van añadiendo registros encadenando nuevos bloques de transacciones formados por x filas del Excel, los cuales, para su formación, dependen de los bloques anteriores. Para añadir los bloques es necesario resolver un problema que requiere un gran poder computacional.

Desde que se creó Bitcoin han aparecido cientos de criptomonedas con características similares, cada una con sus rasgos distintivos. Aun así, a día 14 de febrero de 2020, Bitcoin continúa siendo la criptomoneda más importante representando el 61,47% de la capitalización del mercado de criptomonedas.

Empresas como Facebook, viendo el éxito de las criptomonedas, se han interesado en la tecnología blockchain. Facebook anunció el 18 de junio de 2019 su intención de crear una criptomoneda llamada Libra con el objetivo principal de solventar el problema de que “1700 millones de adultos siguen fuera del sistema financiero al no tener acceso a un banco tradicional, aun cuando 1.000 millones de ellos poseen un teléfono celular y casi 500 millones tienen acceso a internet” (Demirguc-Kunt et al. 2018 citado en “Informe de libra introducción”, s.f., párrafo 3). Esta pretende utilizar la tecnología blockchain para así poder aprovecharse de la gobernanza distribuida, el acceso abierto y una seguridad criptográfica que permita proteger la integridad de los fondos. Tal y como se afirma en la web oficial de Libra, pretende ser una blockchain escalable, aspecto en el que las criptomonedas actuales son deficientes debido a que hay una limitación importante en el número de transacciones que se añade a la cadena de bloques.

Mediante Libra pretenden que haya un nuevo modo de llevar a cabo pagos o transferencias a nivel internacional. Con esto aparece la pregunta sobre las implicaciones que un sistema financiero internacional basado en criptomonedas tendría, que características distintivas tendría respecto al actual, y que problemas podría conllevar en la estabilidad de precios, en la eficiencia de pagos o en las crisis financieras globales que pudieran acaecer.

La segunda parte del trabajo está compuesta de dos partes. En la primera se hace un análisis estadístico de una serie de criptomonedas. Se utilizan los datos del precio de cierre de seis criptomonedas para el periodo comprendido entre el 1 de enero de 2017 y el 31 de diciembre de 2019. Este análisis se hace a partir de la rentabilidad continua, definida como $\ln(p_t / p_{t-1})$, ya que el precio es una variable que posee tendencia y por lo tanto al aplicar la rentabilidad se pretende que la variable transformada sea estacionaria en media y que tenga una varianza constante. Una vez realizada la transformación, se estudia que distribución paramétrica se ajusta mejor a la rentabilidad continua de las principales criptomoneda para así poder analizar posteriormente el riesgo asociado a cada una de ellas con medidas como el CvaR y el VaR.

Además de por los motivos mencionados, se ha escogido esta medida ya que es la más utilizada en casi todos los estudios relativos al análisis estadístico de las criptomonedas cuando se tiene este objetivo. A continuación se muestran los trabajos con un objetivo similar a éste: (Chu, Nadarajah, and Chan 2015) en el cual se ajustan las principales distribuciones de probabilidad a la rentabilidad continua del tipo de cambio del Bitcoin respecto al dólar. Concluyen que la distribución que mejor ajusta es la distribución hiperbólica generalizada; (Chan et al. 2017) en el cual se ajustan las principales distribuciones paramétricas a la rentabilidad continua del tipo de cambio entre diversas criptomonedas y el dólar, concluyendo que no hay una distribución que ajuste mejor en todas y que todas ellas tienen en común un gran apuntamiento y la presencia de asimetría. Por lo tanto, la distribución normal no es una distribución que proporcione un buen ajuste, y en consecuencia, es necesario encontrar otras distribuciones que permitan evaluar el riesgo a través de medidas como son el VaR o el CVaR. En caso de partir del supuesto de normalidad, muy lejos de la realidad, toda conclusión sobre el riesgo quedaría invalidada.

En la última parte del trabajo se estudia la capacidad predictiva de diferentes modelos sobre el precio de cierre de las criptomonedas analizadas en los apartados anteriores. Los modelos utilizados son las redes LSTM, uno de los modelos de Machine Learning más utilizados cuando el objetivo es el de llevar a cabo predicciones a partir de datos secuenciales y en segundo lugar se utiliza el modelo estadístico ARIMA.

Se predice el precio de cierre y no la rentabilidad continua ya que lo que se pretende, una vez se ha ajustado cada uno de los modelos, es analizar los beneficios que se pueden extraer a partir de una estrategia de *trading* muy sencilla. Ésta se hará a partir de las predicciones de los modelos mencionados y servirá como complemento al análisis del riesgo de las criptomonedas hecho con anterioridad. Con esto se consigue tener una idea más amplia de los riesgos y beneficios asociados a la inversión en criptomonedas. Las predicciones de todos los modelos utilizados se compararán entre sí utilizando el error cuadrático medio para así poder concluir que capacidad predictiva posee cada uno de los dos modelos y que diferencias existen entre ellos.

Finalmente se hará una justificación matemática de por qué se ha optado por utilizar las redes LSTM por delante de otras alternativas. La alternativa analizada será la red neuronal recurrente (RNR), modelo de Machine Learning que también sirve para la predicción de datos secuenciales y que servirá además para entender el funcionamiento de las redes LSTM. Las RNR son una simplificación de las redes LSTM y debido a esto tiene varios problemas que ésta última trata de solventar como por ejemplo la incapacidad de preservar la información de momentos del tiempo lejanos para llevar a cabo las predicciones.

2 Metodología

Se han utilizado cuatro bases de datos. Las tres primeras, las cuales han sido utilizadas en la primera parte del trabajo, han sido: “The World Bank, Remittance Prices Worldwide”, que contiene los costes de enviar remesas a nivel internacional, “The Global Findex Database 2017”, que contiene centenares de variables relacionadas con el acceso a servicios financieros y finalmente “The World Bank, Migration and Remittances Data” que contiene datos de la cantidad de remesas enviadas y recibidas por cada país.

La cuarta base de datos es “BNC Digital Currency Indexed EOD”; de ella se obtiene el índice de precios de cierre diario en dólares estadounidenses para todas las criptomonedas estudiadas en el trabajo. Tal y como se describe en la web de Quandl, BraveNewCoin calcula un índice de precios global para cada criptomoneda. Esto se hacía cada 5 minutos hasta el 1 de junio de 2018 y, posteriormente, cada 30 segundos. El índice de precios se basa en datos agregados de todos los intercambios que operan en cada una de las criptomonedas. Para cada mercado de divisas, se obtiene un promedio ponderado por volumen del último precio informado de cada intercambio. Con tal de derivar el índice, cada valor se convierte a dólares estadounidenses utilizando las tasas de conversión fiduciarias internacionales actuales.

Tal y como se ha mencionado en la introducción, la segunda parte del trabajo se compone de un análisis estadístico que sirve para analizar el riesgo asociado a cada criptomoneda y en segundo lugar un análisis en términos de capacidad predictiva mediante el uso de diferentes modelos. Con esto se consigue tener una idea de los beneficios que se pueden obtener invirtiendo en criptomonedas además del riesgo en el que se incurre según en cual se invierta. Por lo tanto, son dos análisis independientes entre ellos pero que se complementan ya que ayudan a disponer información desde tres vertientes: la económica, mediante los beneficios obtenidos con diferentes estrategias de trading, la estadística, mediante el análisis de riesgos y la relativa a los modelos predictivos, permitiendo tener una idea de como de bien se pueden llegar a predecir los precios futuros y como, dependiendo del modelo utilizado, los resultados pueden ser unos u otros.

En la parte estadística la variable estudiada es la rentabilidad continua, que está definida de la siguiente manera:

$$r_i = \ln \left(\frac{P_i}{P_{i-1}} \right)$$

Siendo P_{i-1} el precio de cierre del día anterior al día i y r_i la rentabilidad continua del día i .

Esta es la variable más utilizada cuando se pretende analizar el riesgo asociado a la inversión en activos como pueden ser acciones de una compañía o en criptomonedas como Bitcoin. Esto es debido a que el precio es una variable con tendencia y por lo tanto aplicando esta transformación se consigue que sea estacionaria en media y que la varianza se mantenga constante.

En todo lo relativo a los modelos predictivos se utiliza el precio de cierre ya que se utilizarán distintos modelos como son la red LSTM y el modelo ARIMA para predecir el precio del día siguiente y a partir de estas predicciones poder llevar a cabo una estrategia de trading. Con los

resultados de esta estrategia se evaluarán los beneficios que se pueden llegar a obtener según la criptomoneda y modelo que se escojan.

La estrategia de trading será la siguiente: se parte de un total de 10000 dólares que se invierten en cada criptomoneda. Si se predice que el precio de cierre va a aumentar para el día siguiente se invierte todo en la compra, si se predice que va a bajar, se vende en corto.

La fecha de inicio del 1 de enero de 2017 mencionada en la introducción se ha escogido de tal manera que se posean suficientes datos de las criptomonedas con mayor capitalización (CoinMarketCap 2020) y que estos datos sean lo más actuales posibles. Solo aquellas criptomonedas para las cuales se disponen datos del periodo señalado y además están entre las 15 primeras en capitalización, han sido utilizadas. Estas son: **Bitcoin, Ethereum, Ethereum Classic, Ripple, Litecoin y Monero**. Tether se trata de una criptomoneda estable para la que se disponen los datos pero no se va a analizar su rentabilidad continua ya que el precio es prácticamente fijo.

Se ha ajustado los datos sobre la rentabilidad continua de cada criptomoneda con las principales distribuciones utilizadas en finanzas de las que hay disponibles en el ecosistema SciPy the Python. Estas distribuciones son: t de Student, distribución de Laplace, distribución Normal Inversa Gaussiana, distribución Normal, distribución Logística, distribución exponencial de potencia y distribución normal sesgada. Este ajuste se hace ya que el cálculo del riesgo asociado a cada criptomoneda (mediante medidas como el VaR o CVaR, las cuales sirven para medir el riesgo de un activo o una cartera) se tiene como objetivo hacerlo de forma paramétrica, esto es, a partir de los cuantiles de la distribución final seleccionada. En vez de partir de la asunción de que las rentabilidades continuas siguen una distribución normal, se busca la distribución que mejor los ajusta. La suposición de normalidad es muy común en el mundo de las finanzas y en caso de no cumplirse haría otras metodologías, como la de calcular el riesgo a partir de los datos históricos, una mejor opción. Es por esto que se rechaza esa hipótesis tan restrictiva y se exploran otras mediante el enfoque paramétrico. Una vez se ha seleccionado la distribución que mejor ajusta, se utilizan PP-Plots, QQ-Plots y el test de Kolmogorov Smirnov para asegurarse de que el ajuste es correcto. Un motivo por el que se ha utilizado el enfoque paramétrico es porque se quiere tener un riesgo que tenga mayor capacidad de generalizar a otros periodos ya que el cálculo de este no se ve afectado de una manera tan importante por *outliers*, como es en el caso del enfoque histórico. Ya por último, el uso de esta metodología también tiene la utilidad de ser una fuente de información para aquellos que requieran en su análisis de saber que distribución probabilística es más representativa de la rentabilidad continua.

El criterio para escoger la mejor distribución estará basado en el criterio de AIC. Pese a ello, se muestra el resultado de otros criterios, esto son: Criterio de Información de Bayes (BIC), el Criterio de Información Consistente de Akaike (CAIC), el Criterio de Información Corregido de Akaike (AICc) y finalmente el Criterio de Hannan-Quinn (HQC). Una vez se ha escogido el mejor modelo para cada criptomoneda se usará la prueba estadística de Kolmogórov-Smirnov para determinar la bondad del ajuste. Además, se utilizarán Q-Q Plots y P-P Plots para ver si las distribuciones paramétricas ajustan bien en las colas o en la parte central de los datos. Finalmente, el riesgo se evalúa mediante el VaR y el CVaR, los cuales están calculados de manera paramétrica, a partir de la distribución que proporciona un mejor ajuste.

La base de datos de BraveNewCoin, que contiene los precios de cierre de las criptomonedas, se ha dividido en tres particiones que son:

- *Reduced training set*, que contiene los datos que van del 01/01/2017 al 19/10/2018 (60% de los datos). Estos son los datos con los que se escogerá, tanto para la red LSTM como para el modelo ARIMA, el modelo que proporcione mejores predicciones para cada criptomoneda.
- *Validation set*, que contiene los datos que van del 20/10/2018 al 26/05/2019 (20% de los datos). En estos datos se evaluará la capacidad predictiva de diferentes tipos de modelos ajustados con los datos del reduced training set, con tal de establecer cual es el que permite realizar mejores predicciones.
- *Training set*, que contiene los datos del reduced training set y validation set. Una vez encontrados los modelos que permiten hacer las mejores predicciones se ajustará cada modelo a estos datos.
- *Test set*, que contiene los datos que van del 27/05/2019 al 31/12/2019 (20% de los datos). Con estos datos, los cuales no serán utilizados en todo el trabajo hasta el apartado final, se llevarán a cabo las predicciones con los modelos ajustados con los datos del training set. Con esto se consigue que los modelos finales escogidos, cuya capacidad predictiva final es evaluada en el test set, no hayan sido escogidos siendo influenciados por los datos del test set. En caso de haber escogido el mejor modelo para cada criptomoneda evaluando en el test set se habría acabado teniendo unos resultados que sobreestimarían la capacidad predictiva ya que se acabaría escogiendo el modelo que más capaz ha sido de captar las idiosincrasias del test set.

Se podría haber optado por el uso de *cross validation* pero debido al enorme coste computacional que implicaba se ha optado por utilizar la metodología presentada. Cross validation consiste en dividir el training set en k partes y evaluar los modelos en cada una de las k partes entrenando con las $k-1$ partes restantes. Una vez hecho esto, la capacidad predictiva de cada modelo es el promedio de las k capacidades predictivas. Esta metodología es más robusta que la que se ha utilizado en este trabajo ya que evita que los resultados puedan depender de el periodo escogido para evaluar los modelos.

Los modelos estudiados para la predicción de la variable en nivel (precio), tal y como se ha comentado, son el modelo ARIMA y la red LSTM. Para el ARIMA, el criterio para escoger el mejor modelo ha sido el criterio de AIC (evaluado a partir del ajuste de los datos del reduced training set). A partir de estos modelos se ha evaluado la capacidad predictiva mediante el error cuadrático medio (ECM) en el validation set. No se ha optado por evaluar directamente el ECM en el validation set por razones que se explicarán más adelante y que requerirían de extenderse en demasía. Para la red LSTM, el criterio para escoger el mejor modelo ha sido el ECM, el cual ha sido evaluado directamente en el validation set.

3 Tipos de dinero

En esta primera parte del trabajo, vinculada al grado de economía, se presentan los tipos de dinero que existen o que pueden llegar a existir. Entre los que se encuentran el dinero fiduciario, las criptomonedas, las cuales están basadas en la tecnología Blockchain, las monedas digitales emitidas por Bancos Centrales y las monedas digitales corporativas. Un ejemplo de esta última lo podemos encontrar en Libra, moneda digital que Facebook tiene el objetivo de crear y la cual está teniendo dificultades para ponerse en funcionamiento a nivel mundial debido a las restricciones que ponen muchos países.

3.1 Dinero fiat, vales, E-money

El sistema monetario de casi todos los países a nivel mundial está basado en dinero fiat, en el que el banco central crea dinero de la nada y el gobierno garantiza que el dinero mantiene su valor. El dinero, ha pasado a tener que ser impreso a ser un aumento por parte del banco central en el saldo de las cuentas que los bancos comerciales mantienen en el banco central, sin la necesidad de tener que llevar a cabo una transferencia de dinero para ello (Danielsson 2019).

Antes de continuar hace falta hacer un apunte sobre la diferencia entre dinero fiduciario y el dinero fiat, conceptos que se han visto asociados con muchas contradicciones respecto a su definición en numerosos artículos. Desde que Nixon puso fin al patrón oro a nivel mundial en 1971, el dinero dejó de ser fiduciario y pasó a ser dinero fiat. Ambos tienen la característica de ser dinero que no tiene valor por si mismo. La diferencia está en que el dinero fiduciario no está declarado como una moneda de curso legal, por lo que la población no está requerida por ley a aceptarlo como medio de pago. Además, el emisor del dinero fiduciario promete cambiar este dinero por otro bien (oro de 1944 a 1971) o dinero fiat si se lo piden. Siempre y cuando la población confíe en que la promesa no se romperá, el dinero fiduciario puede usarse como el dinero fiat ("The Four Different Types of Money", 30 de julio de 2019, párrafo 3).

El dinero fiat, por otro lado, no tiene ningún tipo de respaldo y el gobierno lo declara como dinero de curso legal por lo que la población está obligada por ley a aceptarlo como medio de pago. Su valor está derivado de la relación entre la demanda y oferta. Ejemplo de esto se puede encontrar en las monedas y los billetes actuales ("The Four Different Types of Money", 30 de julio de 2019, párrafo 3).

En otras fuentes como ("Dinero fiduciario, economipedia", s.f., párrafo 1-2) se define el dinero fiduciario como toda aquel que está respaldado en la confianza de una sociedad y no está basado en el valor de metales preciosos sino en la creencia general de que el dinero tiene el valor que dice tener. En ("Dinero fiduciario, Wikipedia", s.f., párrafo 5) se afirma lo mismo que en la anterior fuente y además se dice que el dinero fiat es una forma de dinero fiduciario, con la característica distintiva del respaldo legal del Estado.

Los vales son otro tipo de dinero emitido de forma privada y que funciona como una alternativa al dinero. Es una forma de pago que usan algunas empresas para que los empleados los gasten en ciertos servicios como podrían ser los vales de comida.

Por último cabe destacar el e-money, también llamado dinero electrónico, es todo aquel que se

encuentra almacenado como información en un servidor, y esto incluye a las tarjetas de prepago, tarjetas de crédito o los monederos electrónicos. En su concepción más amplia, este dinero se expresa y respalda en unidades monetarias comunes como son el dólar o el euro (“Dinero electrónico, economipedia”, s.f., párrafo 2). Este dinero tiene la característica principal de que no tiene respaldo físico, por lo que tiene la capacidad de fluir como información a través de servidores. Otro ejemplo lo podemos encontrar en PayPal, que permite adquirir todo tipo de productos y servicios en sitios web sin la necesidad de usar tu tarjeta de crédito. Por último, existen otros como AliPay, M-Pesa o WePay.

3.2 Criptomonedas

Las criptomonedas son un medio digital de intercambio que utiliza criptografía para asegurar las transacciones, controlar la creación de unidades adicionales y verificar la transferencia de activos usando tecnologías de registro distribuido (Criptomoneda, s.f.). Ejemplos de éstas son Bitcoin, Ethereum o Ripple. Suelen tener unas características semejantes a las del dinero: son divisibles, fungibles, portátiles y tienen un suministro limitado. Suelen ser utilizadas para el pago de bienes y servicios o como reserva de valor y ahorro, aunque hay excepciones (“¿Cuál es la diferencia entre criptomonedas, tokens, monedas virtuales y digitales?”, 30 de agosto de 2019, párrafo 7).

Las principales diferencias entre el dinero fiat y las criptomonedas son las siguientes (Danielsson 2019):

- Las unidades de dinero son denominadas monedas aunque en el caso de las criptomonedas no tienen una representación física.
- En el caso de las criptomonedas se hace un seguimiento de las monedas en una blockchain y no en una cuenta. La blockchain puede ser permissionada o no permissionada, conceptos se desarrollarán más adelante.
- El dinero fiat, como ya se ha comentado, es creado mediante un aumento en la reserva del banco central, mientras que las criptomonedas son generadas mediante un algoritmo con diferente funcionamiento según de que criptomoneda se trate.

Una vez se han presentado las principales diferencias entre las criptomonedas y el dinero fiat, se procede a explicar la tecnología en la que están basadas las mismas. Esta explicación se hace con el afán de que cuando se explique Libra y las implicaciones financieras de un sistema basado en criptomonedas, se comprenda mejor su funcionamiento y todo lo que hay detrás del mismo, aspecto absolutamente esencial si se pretende entender toda afirmación o conclusión que se saque en este trabajo.

3.2.1 La tecnología blockchain

Tal y como se comenta en (Blockchain basics: Introduction to distributed ledgers, s.f., párrafos 3-4), las criptomonedas registran las transacciones en un libro mayor distribuido el cual es un tipo de base de datos que se comparte, replica y sincroniza entre los miembros de una red descentralizada. Los participantes de esta red gobiernan y acuerdan por consenso sobre las actualizaciones que se hacen el libro mayor. Cada registro en el libro mayor distribuido tiene

una marca de tiempo y una firma criptográfica única, lo que hace que sea un historial auditable e inmutable de todas las transacciones de la red.

Una vez definido el libro mayor distribuido hace falta entender que es una *Blockchain*. La blockchain es un tipo de libro mayor distribuido. Su completo funcionamiento se explicará a partir de imágenes de una Blockchain con transacciones inventadas ya que se considera que mediante una explicación formal no se entendería con claridad todo el funcionamiento, sobre todo si es la primera vez que se presenta este tema. Todas las imágenes que se muestran a continuación se han obtenido en una Demo disponible en ("A web-based demonstration of blockchain concepts - Github, s.f.). Absolutamente todo el contenido de las mismas se irá explicando progresivamente en este apartado. Concretamente, se pondrá el foco en el funcionamiento de la Blockchain de Bitcoin, ya que es la criptomoneda más importante a nivel mundial y el funcionamiento de su Blockchain es parecido al del resto de criptomonedas, aunque cabe destacar que cada una tiene sus características distintivas.

Una Blockchain tiene una estructura muy concreta que consiste en una secuencia de bloques, donde cada bloque está formado por un conjunto de transacciones. Hay unos individuos llamados *miners*, que, mediante la resolución de un problema computacional, consiguen añadir un bloque nuevo de transacciones a la cadena y a cambio son compensados. **Hay que dejar claro desde un principio que lo que se certifica con la tecnología blockchain no es tu balance, sino tus transacciones. Si quieres saber cuanto dinero tienes, necesitas mirar tu histórico de transacciones en la blockchain y hacer los cálculos.** Para entender en que consiste este problema se requiere del entendimiento de un concepto, y este es la función hash.

Las funciones hash, tal y como se definen en (Sánchez de Diego Martínez-Cabrera 2014), son funciones de resumen que utilizan un algoritmo con el que, a partir de una entrada de datos (infinita, cualquier mensaje) producen una salida alfanumérica de longitud fija que resume la entrada, de tal forma que el resultado o valor hash solo puede ser obtenido con los mismos valores del input. Cualquier mínimo cambio en la entrada de datos dará lugar a un output diferente. Por lo tanto, cada input tiene su único hash.

Concretamente, en la figura 1 se muestra la función hash SHA256, utilizada en la Blockchain de Bitcoin, la cual, a partir de un input, que en este caso es la palabra "anders", produce un valor hash de 64 dígitos hexadecimales únicos de 256 bits o 32 bytes que puede verse a la derecha de "Hash:". Esta función Hash no permite decodificar de vuelta, o dicho de otra manera, no permite saber los datos que se han dado como input a partir del output. Ésta es la razón por la cual son tan usados en los algoritmos de encriptación.

SHA256 Hash



Data:	anders
Hash:	19ea4ac2e1a53b1267fe5a61a3b6b81f760ce4223a25b495a5e2b6183da68717

Fig. 1: Ejemplo de aplicación de la función SHA256 Hash. Fuente: A web based demonstration of blockchain concepts, Github

En una Blockchain el input son un conjunto de elementos y el objetivo es encontrar un número, llamado *proof of work*, que junto a esos elementos, dé un valor hash concreto. Este número se dedican a encontrarlo los llamados *miners*, personas que pueden ser particulares que utilizan un poder computacional propio, pueden ser *pools* (particulares que unen su poder computacional y se reparten los beneficios) o empresas dedicadas a esto. Aquel que lo resuelve primero tiene como premio, en el caso de Bitcoin, un número determinado de Bitcoins y una cantidad variable que depende de la tarifa asociada a las transacciones. Además, consiguen que el bloque para el que estaban resolviendo el problema se añada a la cadena. Cada particular o conjunto de particulares decide que transacciones quiere añadir al bloque, dando prioridad a aquellas que tienen mayor tarifa, la cual está definida por aquel que realiza la transacción. Si la transacción le es muy urgente entonces le pondrá una tarifa elevada y así consigue que sea de las primeras en ser añadida al siguiente bloque de la blockchain. Si una transacción no está añadida al bloque no está verificada y por lo tanto es como si no hubiera existido.

En Bitcoin, por ejemplo, tratan de que cada 10 minutos se resuelva uno de estos problemas y por lo tanto que la cadena se vea ampliada cada 10 minutos en un bloque. La dificultad la van modulando para conseguir que esto se cumpla, garantizando que el tiempo como mínimo sea de 10 minutos. Una vez se ha añadido un bloque, el proceso se inicia de nuevo, esta vez trabajando a partir de la Blockchain con el nuevo bloque (siempre la Blockchain con más bloques es la que sirve como referencia y la que se utiliza como cadena a la que añadir nuevos bloques, ya que es la más fiable). Como se puede comprender, si este problema se resuelve cada 10 minutos con centenares de miles de computadores compitiendo entre ellas para resolverlo las primeras, la capacidad computacional que se requiere y el consumo de energía que implican es inconmensurable.

Una vez se ha presentado la función Hash, concretamente la SHA256, que es la que usa Bitcoin, se procede a explicar el funcionamiento de la Blockchain de Bitcoin, que prácticamente igual al del resto de criptomonedas. En la figura 2 se puede ver un ejemplo de una Blockchain formada por cuatro bloques, pero en la que se muestran solo el tercero y el cuarto, el contenido de la cual se explica a continuación:

- El primer elemento, titulado *Block*, hace referencia al bloque de la Blockchain, de tal manera que el bloque número tres tienen dos a la izquierda (bloque 1 y 2, siendo 1 el primer bloque de la Blockchain) y uno a la derecha (bloque 4).

- El *nonce* es un número o proof of work que se ha de encontrar para conseguir que el valor hash (el resultado de aplicar la función hash) tenga un número inicial de ceros determinado, en este caso, a modo de ejemplo, ha de tener cuatro ceros. En la realidad el número de ceros iniciales es bastante superior a esto, haciendo que el esfuerzo computacional requerido sea mucho mayor.
- Tx son el conjunto de transacciones que hay en el bloque, en este caso son 3. El primer elemento de cada fila es el valor en dólares transferidos en la criptomoneda de la Blockchain, el segundo elemento es el nombre de la persona que envía el dinero, y el tercero el nombre de la persona que lo recibe.
- Prev es el valor hash resultante del anterior bloque.
- Hash es el valor hash de este bloque. Para el bloque 4 está calculado mediante la función SHA256 de la siguiente manera:

$$SHA256(Proof\ of\ work, transacciones, valor\ Hash\ bloque\ previo) =$$

$$= SHA256(116068, las\ cinco\ transacciones, 0000a9dd50de891b2...) = 0000aa5cceedd53...$$

Block: # 3

Nonce: 40596

Tx:

\$	3.14	From:	rester	->	Bugs
\$	2.12	From:	Twee	->	Roadr
\$	1.99	From:	Daffy	->	Marvi

Prev: 0000f843c73a7b3f5f3af6b7a4f5690a377326

Hash: 0000a9dd50de891b2de8601c6d933c586152

Mine

Block: # 4

Nonce: 116068

Tx:

\$	52.19	From:	Rick	->	Ilsa
\$	867.5	From:	Capt	->	Stras
\$	276.1	From:	Victor	->	Ilsa
\$	97.13	From:	Rick	->	Sam
\$	119.6	From:	Capt	->	Jan Br

Prev: 0000a9dd50de891b2de8601c6d933c586152

Hash: 0000aa5cceedd53f9078325617d14f0c28903

Mine

Fig. 2: Blockchain formada por cuatro bloques en la que se muestran los dos últimos. Contiene un conjunto de transacciones inventadas que sirven de ejemplo. Fuente: A web based demonstration of blockchain concepts, Github

Esta función SHA256, con estos inputs en concreto, da como resultado el valor hash que cumple con la condición de tener los cuatro primeros elementos como 0, por lo tanto, se ha resuelto correctamente el problema y es por esto que este bloque 4 se ha podido añadir a la Blockchain. En conclusión, dadas unas transacciones y el hash del anterior bloque, lo que hay que ir cambiando

es el valor que está definido como “Nonce” en la imagen hasta conseguir que el valor hash del bloque que se está tratando de añadir tenga cuatro ceros iniciales.

Como puede observarse, el bloque 4 tiene como Prev el Hash del bloque 3, el cual le sirve como input junto a las cinco transacciones y el valor “Nonce” o *proof of work* (116068) que hay que encontrar. Esto garantiza la integridad de la Blockchain, ya que si se quiere modificar una transacción del bloque 3 de forma fraudulenta, suponiendo que ya se había añadido el bloque 4 a la Blockchain, esto afecta al proof of work del bloque 3 (pasaría de ser 40596 a otro número debido a que las transacciones han cambiado) y consecuentemente afecta al hash del bloque 3 que a su vez afecta al hash del bloque 4 a causa de que éste lo usa como input. De esta forma, si lo que se quiere es alargar la cadena añadiendo el bloque 5 y a su vez modificando fraudulentamente el bloque 3, hay que resolver el proof of work tres veces consecutivas antes que el resto del mundo, cosa que es prácticamente imposible. Es por esto que se le llama Blockchain o cadena de bloques; porque los bloques están encadenados a través del valor hash del bloque previo.

El lector podría estar preguntándose como se garantiza primero de todo que la persona que realiza la transacción tiene los tokens (unidades de la criptomoneda que se está usando en la transacción) que dice tener y en segundo lugar como se garantiza que una persona no pueda añadir una transacción a la Blockchain haciéndose pasar por ti. Aquí es donde entra la clave privada y la pública. En la figura 3 se puede ver un ejemplo, donde dada una clave privada que eliges al azar y la cual solo tú conoces, generas una clave pública de tal manera que solo tu clave privada genera esa clave pública. A partir de la clave pública no se puede obtener la clave privada, pero sí que se puede hacer el proceso inverso. Por lo tanto, la clave pública todo el mundo la puede ver cuando realizas una transacción mientras que la clave privada no.

The image shows a web interface for generating key pairs. It has a title 'Public / Private Key Pairs'. Below the title, there are two input fields. The first is labeled 'Private Key' and contains the text '65645124001427200769618650116875726431131717237126312693843499695125187055170'. To the right of this field is a button labeled 'Random'. The second field is labeled 'Public Key' and contains the text '0433d43659db98f64ea8809ee0e4a8a2a28882b95cc3f852505adc88b738ec8af49570651186759421fa554ff5f8bcacc208bedb19e8f'. The fields are light blue and the text is black.

Fig. 3: Ejemplo de creación de la llave pública a partir de la llave privada. Fuente: A web based demonstration of blockchain concepts, Github

Cuando alguien quiere realizar una transacción tiene que generar una firma, como es el caso de la figura 4. En esta la firma o *signature* se calcula mediante una función f definida de la siguiente manera:

$$\begin{aligned} f(\text{Mensaje}, \text{llave privada}) &= \text{Firma} = \\ &= f(\text{Hello Anders}, 0433d43659...) = 3045022100dbe... \end{aligned}$$

El Mensaje es lo que define que transacción quieres realizar y la llave privada es la que es propiedad del que la realiza. En esta figura se ha puesto una frase al azar. Antes de añadir una

transacción se ha de hacer una verificación de que esta es real y ha sido realizada por la persona propietaria de la llave pública. Esto se hace mediante la función v , la cual da como output “TRUE” si la firma ha sido realizada por la persona que posee esa llave pública (se demuestra que conoce la clave privada correspondiente a esa clave pública) y “FALSE” si no es el caso. Solo si la salida es TRUE se puede dar como válida la transacción. La función V es la siguiente:

$$V(\text{Mensaje}, \text{Firma}, \text{Llave pblica}) = T/F =$$

$$V(\text{Hello Anders}, 3045022100dbe..., \text{Llave } 0433d43659...) = T/F$$

The screenshot shows a web interface titled "Signatures". It has two tabs: "Sign" (selected) and "Verify". Under the "Sign" tab, there is a "Message" input field containing "Hello Anders". Below the message is a "Public Key" input field containing a long hexadecimal string: "0433d43659db98f64ea8809ee0e4a8a2a28882b95cc3f852505adc88b738ec8af49570651186759421fa554ff5f8bcacc208bedb19e8". Below the public key is a "Signature" input field containing another long hexadecimal string: "3045022100dbe782f3de621bd30fb87fe5150bd549fcf7c9e3c35cb9466b622fd25d95a60f022038b80f9c38a6be399cf705488ee670". At the bottom of the interface is a large blue button labeled "Verify".

Fig. 4: Creación de la firma de una transacción a partir del mensaje y la llave pública del que la realiza. Fuente: A web based demonstration of blockchain concepts, Github

Una vez ya se ha entendido todo esto se va un paso más allá y en la figura 5 se puede ver como sería una hipotética blockchain. En esta hay un elemento extra que no se ha explicado que es la Coinbase. Para simplificar, en esta blockchain se supone que solo hay un individuo poseedor de tokens. En la propia Blockchain, que se ha supuesto que es pública como en el caso de Bitcoin y por lo tanto todo aquel que quiera, puede acceder a ella, mirando en bloques pasados, se puede ver si una persona realmente tiene los tokens que dice tener, ya que si en ningún momento ha recibido, entonces es imposible que pueda enviar. De esta forma se garantiza que todas las transacciones sean viables. La Coinbase indica cuantos tokens ha recibido un individuo concreto y por lo tanto te permite hacer esta verificación.

Ahora las transacciones se muestran como se ven en la realidad. En estas se ve la cantidad enviada, quién la envía (su llave pública), la llave pública del que recibe los tokens y la firma de la transacción. Todas estas transacciones han sido verificadas a través de la función V y por lo tanto se ha garantizado que el que las realiza es quién dice ser.

The image displays two panels of a blockchain simulation interface, representing two consecutive blocks in a chain.

Block # 4 (Left Panel):

- Block:** # 4
- Nonce:** 63022
- Coinbase:** \$ 100.00 → 04fe1be031bc7a54d900ff062911b
- Tx:**
 - \$ 15.00 From: 04d4080959e3795b → 0451d4a9c44a2dec Sig: 3045022100fd0c2534ba9c1c3f947e4d29ac5f54442ce9e03f3dc8dd285260
 - \$ 5.00 From: 04222d7af343abd → 041c377677bb6973 Sig: 304402200b8d07fe4949a8eb958262d1fe579a5f0f96c2b4e1aa97a41ae0102
 - \$ 8.00 From: 04cc17dc129331c1 → 04d4080959e3795b Sig: 30440220665c64c85982f75d78aa9957a6a805ed4999f8ad183d4cea7f7c507
- Prev:** 0000a9e2a5d6100c1fa23580671cc4f3bca3c58180d0f55e49f9e49a389f2777
- Hash:** 0000e0e3d78d093313f15936fb3d08f06b2bd095044342a1c896a3ee8b10a7bf
- Mine** button

Block # 5 (Right Panel):

- Block:** # 5
- Nonce:** 7355
- Coinbase:** \$ 100.00 → 04cc17dc129331c1cbb9c32cf4dc2
- Tx:**
 - \$ 2.00 From: 04d4080959e3795b → 04d84dae793a8253 Sig: 304502207fcc9d79c7894a4fa246f3b1ee8b21a40ae7f195e8f08ffe253d163
 - \$ 6.00 From: 0451d4a9c44a2dec → 043e17e5095e878b Sig: 30450220454632e38948141be2c1b75e6c08b2b98dfc6e95d1691cdd6cd0a31
 - \$ 4.00 From: 0451d4a9c44a2dec → 04020d6fe7aeabd3 Sig: 3046022100e5e8cb0d2ac042cc8c026c52622a191780da1bdca41ebe2b6190f
 - \$ 9.95 From: 040b4c84f02bfec4 → 04148850d1edbd66 Sig: 304502203f18249ae65e941f0571cc58deb3455700f2508c6ad04ba45194a6
- Prev:** 0000e0e3d78d093313f15936fb3d08f06b2bd095044342a1c896a3ee8b10a7bf
- Hash:** 000056aeaa2c748970c80ac2412be73e0cec4593862b710c59d0d3f90764cd
- Mine** button

Fig. 5: Blockchain formada por cuatro bloques. Fuente: A web based demonstration of blockchain concepts, Github

3.2.2 Tipos de Blockchain

Aunque hay varios tipos de blockchain, se presentarán las dos principales, que son la blockchain pública y la privada o permissionada:

- **Blockchain pública:** es la blockchain en la que están estructuradas casi todas las criptomonedas, incluido Bitcoin. Permiten el acceso a cualquier persona, los usuarios son anónimos, se suelen ofrecer recompensas por el minado de los bloques por parte de los usuarios y la base de datos con todas las transacciones puede ser mantenida por todos los usuarios que lo deseen. Ésta última es una característica muy común de las criptomonedas, ya que tienen un alto nivel de descentralización debido a que quien quiera puede tener la blockchain en su ordenador, garantizando así que haya más seguridad de que no se corrompa. Como no hay ningún organismo regulador, este tipo de blockchain se basa en el consenso de los usuarios de la misma, los cuales toman como referencia la blockchain con más bloques como la real.
- **Blockchain privada o permissionada:** este tipo de blockchain rompe con la idea original de tener una base de datos abierta y accesible a todo aquel que quiera. En estas se requiere permiso para acceder. El control no es ejercido por la comunidad que las usa sino por una entidad centralizada que se encarga de mantener la cadena y aceptar los bloques. Por lo tanto, en éstas no existe la actividad del mining. Por último, cabe destacar que el dueño o dueños controlan que información puede ver el público de la misma y controlan todos los aspectos relativos a ella, incluida su estructura y cambios que puedan producirse.

Un rasgo distintivo muy importante entre estas dos es que mientras que con la blockchain pública las transacciones suelen ser validadas a un ritmo muy lento, con la blockchain privada pueden validarse a un ritmo mucho más elevado. Esto es algo que será analizado más adelante.

Es por razones como esta que ha aparecido una alternativa a las criptomonedas, con unas ventajas que estas no tienen, llamada Libra, la cual estará basada en la tecnología blockchain, pero será permissionada. Entra dentro de lo que se llaman monedas digitales corporativas, las cuales se desarrollan a continuación.

3.2.3 Diferencia entre criptomonedas, altcoins, stablecoins, tokens, monedas virtuales y monedas digitales

A lo largo de la investigación realizada para escribir este trabajo se ha encontrado una enorme confusión respecto a estos conceptos, llegando a ver errores graves de concepción en páginas web con mucha relevancia a nivel nacional e internacional. Es por ello que se opta por dejar claras cuales son las diferencias para que el lector lo tenga todo claro.

En términos generales, la diferencia entre una criptomoneda y un token es que las criptomonedas utilizan su propia blockchain y los tokens utilizan la blockchain de una criptomoneda para funcionar. Entre criptomonedas y tokens hay 2368, aunque los segundos son los más comunes. A diferencia de éstas, los tokens tienen una función diferente a la de ser dinero, ya que estos pueden llegar a representar acciones o pueden ser canjeados por servicios (“Criptomonedas: ventajas, riesgos y futuro según tres expertos”, 27 de setiembre de 2019, párrafo 34).

Token no es más que es un nuevo término para una unidad de valor emitida por una entidad privada, tiene un valor que es aceptado por la comunidad y se fundamenta en la tecnología blockchain. Es parecido a bitcoin pero con la diferencia que tiene muchos más usos. La mayoría de estos se asientan sobre el protocolo de blockchain de Etheruem (“Qué es un ‘token’ y para qué sirve”, 7 de julio de 2017)

Ethereum es una red descentralizada de computadoras con dos funciones principales (“Tokens ERC20: ¿Qué son y cómo funcionan?”, párrafos 4-5):

- una blockchain que permite registrar transacciones
- una maquina virtual que permite la creación de contratos inteligentes

Es capaz de soportar lo que se denomina como aplicaciones descentralizadas llamadas DApps, las cuales están construidas sobre la cadena de bloques Ethereum y aprovechan su tecnología. Fue creada con la idea de acabar con el intermediario que se encuentra en la mayoría de operaciones hoy en día.

Todas las apps que forman parte de Ethereum comparten una misma blockchain en la que pueden conservar la información que quieran y en la cual hay un lenguaje de programación y moneda que todas comparten. La pieza fundacional detrás de estas DApps están los contratos inteligentes. Para entender esto se pone un ejemplo.

En la situación de que un granjero de plátanos de España quiere tener un seguro en caso de que haya una sequía. Antes, su única opción era ir a una compañía aseguradora y tomar una póliza

de seguros, pagando por ello x dinero a la aseguradora y recibiendo y si sufre de un incendio. Además, el granjero tiene que confiar plenamente en todos los intermediarios como son las cortes y la aseguradora.

Mediante un contrato inteligente de seguro todo esto es muy distinto. Este contrato contendría un código que diría: “transfiere x dólares del granjero, en caso de que durante t días no llueva, paga en este caso y al granjero”. Estos contratos suelen ser públicos (cualquiera puede tener acceso a ellos) y además se aplican por si solos si se cumplen las condiciones establecidas en su código por lo que esto da las garantías necesarias de su cumplimiento. Estos contratos, los cuales son registrados en la blockchain de Ethereum, utilizan las criptomonedas llamadas Ether, que pueden ser utilizadas como formas de pago e inversión o para los contratos inteligentes. La ventaja que tienen es que una compañía puede crear cientos de contratos inteligentes que se ejecutan de forma automática en caso de cumplirse ciertas condiciones escritas en su código.

Algunas criptomonedas que utilizan la tecnología blockchain, entre ellas Ethereum, permiten la creación de tokens. Por ejemplo el token ERC-20, los cuales son diseñados y utilizados en la plataforma de Ethereum, permiten que las apps puedan crear fichas ERC-20 las cuales pueden funcionar como una moneda, una acción en la compañía, como prueba de propiedad de una cierta cantidad de oro o de la escritura de una casa, etc. (“Tokens ERC20: ¿Qué son y cómo funcionan?”, párrafos 4-5). ERC-20 es el lenguaje universal que utilizan todos los tokens de la red Ethereum que permiten su negociación.

Los contratos inteligentes en Ethereum son los responsables de la creación de estos tokens, de facilitar las transacciones con estos y de registrarlos en una cuenta. Un ejemplo real de token es el de BAT, token que se usa dentro de las DApps, que entra de los denominados “tokens de utilidad”, que usan la plataforma blockchain de Ethereum. Los anunciantes compran anuncios con estos tokens BAT y luego estos asunciones se distribuyen entre los usuarios en un navegador web llamado Brave sin tener que depender de un intermediario.

Los altcoins son todas las criptomonedas que vienen después de Bitcoin, algunas de las cuales son una bifurcación de la misma como Litecoin o Dogecoin y otras que están construidas en una blockchain nueva como Ethereum. Para poder llamarla altcoin la criptomoneda tiene que estar basada en su propia blockchain. La altcoin más famosa que existe es Ethereum.

Los *stablecoins* son criptomonedas que mantienen un tipo de cambio constante con monedas fiat. Ejemplo de esto se puede encontrar en Tether.

Las monedas virtuales, fueron definidas por el Banco Central Europeo en 2012 (“¿Cuál es la diferencia entre criptomonedas, tokens, monedas virtuales y digitales?”, 30 de agosto de 2019, párrafo 16) como “dinero digital en un entorno no regulado, emitido y controlado por sus desarrolladores y utilizado como método de pago entre los miembros de una comunidad virtual específica”. Un ejemplo claro de esto son el dinero que presente en muchos videojuegos.

Las monedas digitales se definen como toda forma de dinero electrónico, ya sea una moneda virtual, una criptomoneda o un token. Solo existen en forma digital y electrónica a diferencia de los billetes y monedas que usamos en el día a día (“¿Cuál es la diferencia entre criptomonedas, tokens, monedas virtuales y digitales?”, 30 de agosto de 2019, párrafos 14-15).

3.3 Monedas digitales corporativas: Libra, concepción original

Este tipo de monedas tienen mucho en común con los vales emitidos por las empresas, ya que tienen la función de ser una alternativa al dinero y a la vez tiene componentes de los sistemas de e-money ya que por ejemplo, en el caso de Libra, moneda digital que quiere crear Facebook, se podrá conseguir realizar pagos o transacciones desde aparatos electrónicos como podría ser el móvil. Tal y como se menciona en (Danielsson 2019), no se puede considerar una criptomoneda pese a los rasgos de parecido que esta tiene con las mismas debido al hecho de que, para ser una criptomoneda, ha de existir en una blockchain pública a la que todo aquel que quiera tiene acceso, pero ese no es el caso. Pese a esto, se suele hablar de ella como una criptomoneda, aunque en este trabajo, debido a lo argumentado en este párrafo, se considerará como una moneda digital corporativa.

Facebook, con esta moneda digital, tiene el objetivo de solventar el problema de que “1700 millones de adultos siguen fuera del sistema financiero al no tener acceso a un banco tradicional, aun cuando 1.000 millones de ellos poseen un teléfono celular y casi 500 millones tienen acceso a internet” (Demirguc-Kunt et al. 2018 citado en “Informe de libra introducción”, s.f., párrafo 4). Además, hay otros aspectos en los que quiere tener incidencia, y es en los altos costes asociados a los servicios financieros o el acceso limitado a servicios financieros que hay en algunos países subdesarrollados, sobre todo en África.

Tal y como se menciona en (“Consumer Federation of America. How Payday Loans Work”, citado en “Informe de libra introducción”, s.f., párrafo 5), “los préstamos de adelanto de sueldo pueden llegar a tener tasas de interés anualizadas del 400% o más y cargos financieros de hasta 30 dólares por préstamos de solo 100”. Se preguntó a las personas que están al margen del sistema financiero actual el motivo por el que están en esa situación y estos respondieron que es debido a que “carecen de fondos suficientes, las comisiones son altas e impredecibles, los bancos están demasiado lejos, y no cuentan con la documentación necesaria.” (Demirguc-Kunt et al. 2017 citado en “Informe de libra introducción”, s.f., párrafo 5).

Ante la pregunta de que por qué Libra es necesaria si ya existen criptomonedas con las que se pueden realizar transacciones, la respuesta que dan es que no hay una adopción generalizada de los sistemas de blockchain, las criptomonedas presentan mucha volatilidad y falta de escalabilidad. Además, son principalmente reservas de valor o medios de cambio que presentan muchos problemas debido a la capacidad que tienen de evadir regulaciones y a las escasas medidas que se han tomado para evitar el lavado de dinero (“Informe de libra introducción”, s.f., párrafo 7). Es debido a esto que Libra tiene las siguientes características distintivas (“Informe de libra introducción, Sección 2”, s.f., párrafo 4):

- Tiene una blockchain segura, escalable y confiable.
- Tiene el respaldo de una reserva de activos para que Libra tenga un valor intrínseco, a diferencia que lo que pasa con todas las criptomonedas, cuyo valor está basado en el estado del mercado.
- Está regida por el gobierno de la llamada Asociación Libra, que tiene un carácter independiente.

Por lo tanto, un rasgo distintivo de Libra es que está mantenida por una red distribuida de nodos validadores, los cuales son los miembros de la Asociación Libra. En consecuencia, se trata de una blockchain privada con escaso nivel de descentralización.

Para formar parte de la Asociación Libra se tiene que realizar un pago mínimo de 10 millones de dólares (“informe de libra”, s.f.). El objetivo de Libra es llegar a los 100 miembros, pero tal y como se verá más adelante y como se menciona en (“Here’s what you need to know about Facebook’s controversial Libra cryptocurrency”, s.f.), algunos de los miembros fundadores como PayPal, eBay, Visa o Mastercard, clave por su entendimiento en todo lo relacionado con los pagos y la tecnología de transferencia de dinero, han abandonado la asociación.

Continuando con los aspectos diferenciadores de Libra respecto a las criptomonedas convencionales, hay un aspecto a destacar y es la descentralización. Mientras que en criptomonedas como Bitcoin decenas de miles de personas que tienen la blockchain en sus ordenadores, en el caso de Facebook se tiene pensado que esté en manos de 100 miembros aunque se tiene la expectativa de que en el futuro este número sea mayor. Por lo tanto, Libra es descentralizada, pero no en las dimensiones de otras criptomonedas.

Relativo a la resistencia, censura o bloqueo, mientras que Bitcoin no tiene, ya que no está sujeta a casi ninguna regulación, en el caso de Facebook, si se quiere implementar en un país o agrupación de países como es la Unión Europea, tiene que cumplir con muchas regulaciones y en algunos casos con el rechazo absoluto debido al impacto económico que podría llegar a tener.

Finalmente cabe destacar el valor de las monedas. Mientras que Bitcoin destaca por su escasez, aspecto que explica en parte porque tiene un valor de mercado tan elevado, Libra quiere ser una moneda digital estable y vinculada a un conjunto de monedas fiat para añadirle la estabilidad que las criptomonedas no poseen y así conseguir que sea un medio adecuado para la realización de transacciones.

3.3.1 La Reserva de Libra

3.3.1.1 Funcionamiento y estructura

Como se ha comentado en el apartado anterior, el valor de Libra se vinculará a una cesta de monedas fiat y tendrá el respaldo de una reserva de activos. Este segundo elemento es el que se va a desarrollar en este apartado, ya que del primero solo cabe decir que la vinculación se hará con las monedas con más relevancia a nivel mundial como son el dólar o el euro. Por lo tanto, el valor de Libra dependerá de la evolución que tengan las divisas que forman parte de la cesta.

Con tal de dar confianza en su valor intrínseco, la Libra tendrá el respaldo de una cesta consistente en (“La Reserva de Libra”, s.f., párrafo 9) :

- Depósitos bancarios
- Títulos gubernamentales a corto plazo (plazo inferior a 18 meses) de bancos centrales con buena reputación, que se consideran una inversión financiera muy segura debido al hecho de que, por un lado, cuentan con el respaldo del Gobierno, y por otro lado, tienen un vencimiento corto. De esta forma Libra consigue solucionar el problema de la enorme volatilidad que presentan las criptomonedas y permiten que cualquier usuario que quiera

vender cualquier moneda Libra lo que conseguirá hacer a un valor muy próximo al de la reserva. Esta reserva solo puede incrementarse de dos maneras, las cuales garantizan su valor intrínseco (“La Reserva de Libra”, s.f., párrafo 9):

- Aportaciones de los miembros fundadores, los cuales recibirán un incentivo a cambio.
- Los usuarios cuando realicen una compra de Libra mediante el equivalente en dinero fiat.

Estas aportaciones de los usuarios no harán que reciban ningún retorno, ya que la reserva invertirá con ese dinero en activos de bajo riesgo como puede ser deuda de Gobiernos estables con una probabilidad de incumplimiento baja. El interés generado se destinará a cubrir gastos operativos como podrían ser inversiones de crecimiento y desarrollo del ecosistema o investigaciones. Para garantizar que Libra sea capaz de ajustarse a cambios bruscos de demanda, estas inversiones solo se harán en títulos de corto plazo cotizados en mercados líquidos con un alto nivel de volumen de transacciones diario (“La Reserva de Libra”, s.f., párrafo 9).

A cargo de la conservación de la reserva estará una “red global de guardianes” distribuidos en zonas del mundo muy separadas para conseguir salvaguardar los activos presentes en la reserva, garantizar unos altos niveles de transparencia y auditabilidad y además evitar tener todos los riesgos asociados a tener una reserva centralizada (“La Reserva de Libra”, s.f., párrafo 8).

3.3.1.2 Interacción entre las entidades y la reserva

Hasta ahora se ha comentado como será estructura de la reserva pero no se ha hablado de como se hará posible que los ciudadanos intercambien dinero fiat por Libra. Para conseguir esto los usuarios no tendrán una interacción directa con la reserva, sino que una serie de revendedores autorizados (casas de cambio y bancos), que serán las únicas entidades con autorización para llevar a cabo transacciones de cantidades grandes de fiat y Libra desde y hacia la reserva. Para conseguir aumentar las posibilidades de compra de Libra por parte de usuarios estos revendedores se integrarán con plataformas de cambio e instituciones que se dedican a la compra y venta de criptomonedas a usuarios (“La Reserva de Libra”, s.f., párrafo 10). Para garantizar que Libra no puede depreciarse ni apreciarse, la quema o acuñación de las monedas va a depender de la demanda de los revendedores autorizados, de tal manera, que para acuñar monedas, los revendedores autorizados deberán realizar un pago equivalente en activos fiat a la reserva (“La Reserva de Libra”, s.f., párrafo 11).

Para aumentar aún más el uso de Libra y hacerlo a llegar a personas que quizá no tienen los medios o conocimiento para intercambiar en plataformas online, Libra está considerando tener relaciones con empresas comerciales e instituciones reconocidas para así conseguir que haya más revendedores autorizados y permitir que las personas puedan cambiar su moneda local con más facilidad.

3.3.1.3 El cambio de la reserva a lo largo del tiempo

La canasta de monedas a las que se vinculará Libra puede sufrir cambios si hay una crisis económica. Cualquier modificación de tal relevancia como esta requeriría del voto por mayoría cualificada del consejo de la asociación. Salvo en situaciones como ésta, siempre se tratará de que Libra se venda con un estrecho margen por encima y debajo del valor de reserva subyacente; para conseguir esto tiene que haber un mercado competitivo como el mencionado en el apartado anterior (muchas entidades intermediarias) para así evitar que algunas entidades impongan tasas de cambio abusivas (“La Reserva de Libra”, s.f., párrafo 15).

Para evitar retiradas masivas de la moneda en caso de que haya una crisis global o local, como el valor de mercado de la reserva siempre respalda el valor de Libra, no existe el incentivo de retirar las monedas Libra por miedo a que el valor subyacente sea inferior al nominal o que el valor de Libra baje. Gracias a un entorno competitivo como el que se ha comentado, pretende conseguir que hasta en estas situaciones extremas puedas cambiar Libras por dinero fiat con un valor lo más próximo a su valor actual (“La Reserva de Libra”, s.f., párrafo 16).

3.3.1.4 Política monetaria

La asociación decidió que no crearía su propia política monetaria, sino que heredaría las políticas de los bancos centrales que tengan representación en la canasta. Tal y como se menciona en (“La Reserva de Libra”, s.f., párrafo 17), el funcionamiento será parecido al de las cajas de conversión.

Los países que logran una estabilización del nivel de precios tienen un factor en común que es el de tener una emisión de dinero compatible con la demanda de dinero. No se entra en el debate eterno de si la política monetaria (PM) tiene efectos reales a largo plazo en la economía (el dinero no es neutral a largo plazo), visión que encaja con la de Keynes o si la PM solo provoca inflación a largo plazo (el dinero es neutral a largo plazo), que es la visión monetarista.

Uno de los primeros mecanismos para tratar de regular los precios y la balanza comercial fue el patrón oro, que estuvo presente entre 1870 y 1913, bajo el cual todo país garantizaba que su moneda podía ser totalmente convertible a oro. Si los ciudadanos demandaban al banco central llevar a cabo esta conversión, éste tenía la obligación de hacerla, consiguiendo así que los países no pudieran aumentar la oferta monetaria de forma descontrolada, ya que tenían que garantizar la convertibilidad. No fue sino hasta 1913 que los países europeos, dado el efecto que la guerra tenía sobre sus economías, que empezaron a imprimir billetes muy por encima de lo que las reservas de oro podían respaldar, que el sistema llegó a su fin.

Otro sistema que se adoptó fue el que surgió de los acuerdos de Bretton Woods en 1944, a partir del cual se creó el Banco Mundial y el Fondo Monetario Internacional y además se estableció como moneda de referencia internacional el dólar. Estados Unidos tenía que mantener el precio del oro en 35 dólares por onza y tenía la capacidad de cambiar dólares por oro a ese precio sin limitaciones. El resto de países tenían la obligación de fijar el precio de su moneda en relación con el dólar, teniendo que intervenir en caso de necesidad para mantener el tipo de cambio dentro de una banda de fluctuación del 1% (“Acuerdos de Bretton Woods”, s.f.). Este sistema acabó en 1970 dada la desconfianza creciente sobre la convertibilidad del dólar al oro y debido a que países como Francia demandaron grandes cantidades de oro a cambio de dólares.

Es en este contexto que surgen otras alternativas como las cajas de conversión, que es parecido a lo que quiere implementar Libra, por la mecánica de interacción que tiene con la reserva. Éstas se han aplicado en Hong Kong (1983), Argentina (1991) y Estonia, siendo exitosas. En Estonia permitió acabar con la hiperinflación que hubo hasta 1993 así lograr un 20\$ anual a comienzos de 1995, en el caso de Argentina pasando de tener hiperinflación en 1991 a tener unas tasas próximas al 3% en 1994 (“Caia de conversión”, 1995, párrafo 7).

Con este sistema, tal y como se comenta en (“La Reserva de Libra”, s.f., párrafo 17), “las cajas de conversión solo imprimen moneda local cuando existen suficientes activos en moneda extranjera para respaldar por completo una nueva acuñación de billetes y monedas.”. De esta forma, lo que se consigue es que la emisión de dinero corresponda con la demanda de dinero, ya que el dinero emitido no puede ser superior al nivel de reservas internacionales. De esta forma se crea una limitación a la cantidad de crédito que un banco central puede crear para financiar al gobierno o el sector privado.

A diferencia de lo que pasa con el tipo de cambio fijo, con las cajas de conversión siempre hay una garantía de que el banco central va a tener reservas internacionales equivalentes a todo el dinero en circulación, mientras que con el cambio fijo en gran cantidad de ocasiones los bancos centrales han emitido dinero muy por encima de las reservas que tenían, ya sea para la expansión del crédito doméstico o para financiar al gobierno (“Caia de conversión”, 1995, párrafo 6).

En conclusión, Libra adopta este sistema por dos motivos (“La Reserva de Libra”, s.f., párrafo 17):

- La estabilidad debida a que los activos subyacentes seleccionados tienen baja volatilidad.
- La protección contra usos indebidos que podrían haberse dado en otras circunstancias como son la impresión de monedas adicionales sin un respaldo existente, aspecto que hubiera generado desconfianza en Libra.

3.4 Monedas digitales de bancos centrales

La mayoría de bancos centrales están estudiando la posibilidad de emitir su propia moneda digital llamada Central Bank Digital Currency (CBDC) debido al miedo que hay a que entidades privadas puedan crear sistemas de pago que estén fuera del control de los gobiernos (Danielsson 2019). Este mismo artículo menciona que los Bancos Centrales podrían estar interesados en un sistema basado en CBDC, y da las razones siguientes:

- En la actualidad hay países que se están deshaciendo del dinero. En [“Sweden: how cash became more trouble than it’s worth”, 9 de marzo de 2019, párrafos 1-5], se menciona como, en el año 2018, solo uno de cada 8 suecos usaron dinero en efectivo para una compra. Allí hay establecido un sistema de pagos llamado Swish que te permite hacer transferencias instantáneas de dinero usando únicamente tu teléfono móvil, el cual puede ser utilizado para cualquier transacción. Ante situaciones como esta o parecidas, como es el caso de Libra, los Bancos Centrales tienen miedo de perder el control de los sistemas de pago y los reguladores temen no poder asegurar la protección del consumidor ante situaciones como por ejemplo el blanqueo de dinero. Es por esto que la opción de emitir su propia moneda digital para neutralizar esta posible amenaza es una alternativa a estudiar.

- Este sistema puede ser más eficiente y permitir pagos más rápidos y con más seguridad.
- Las CBDC, al no ser anónimas como se explicará a continuación, dan más control al estado sobre la población.
- El banco central tiene más facilidad para controlar la oferta de dinero. A día de hoy si el banco central quiere cambiar el nivel de precios o estimular la inversión en la economía tiene la opción de hacer expansión cuantitativa u operaciones de mercado abierto donde ambas implican la compra o venta de activos financieros, siendo la primera una opción igual que la segunda solo que a una escala mucho más grande y donde en vez de solo comprar deuda pública del estado a corto plazo, compra todo tipo de activos financieros para estimular todos los sectores de la economía. También puede reducir el tipo de interés salvo que se esté en la situación de que el tipo de interés sea 0 o muy próximo a cero. El aspecto positivo de usar CBDC es que da la opción al banco central de afectar la oferta y demanda de esta moneda digital de una forma inmediata cambiando el valor de la misma, sin necesidad de tener que recurrir a las operaciones de mercado abierto.

La CBDC tiene la intención de servir como moneda de curso legal, y estar disponible y accesible para cualquier residente del país, aunque tal y como se menciona en (Griffoli et al. 2018), para que esto sucediera se necesitan cambios en la legislación, ya que la definición de curso legal, aplicado normalmente a monedas y billetes emitidos por los bancos centrales, varía según las jurisdicciones y de hecho, en algunas no hay una obligación por parte del creditor de aceptar una moneda corriente como método de pago. Las CBDC's tendrían una amplia disponibilidad para pagos minoristas (transacciones entre dos consumidores, entre consumidores y empresa y entre empresas) y las reservas, que pueden considerarse como una forma mayorista de CBDC pensada exclusivamente para pagos interbancarios, serían utilizadas sobre todo por bancos con cuentas en el banco central. A diferencia del dinero de curso legal de hoy en día, la CBDC no tiene la intención de tener una forma física.

En (Griffoli et al. 2018) aclara que los efectos que la CBDC podría tener fuera de la frontera no se han considerado aún. De hecho es algo que cabe estudiar con profundidad, ya que podría darse la situación de que se requiriese un tipo de cambio entre las diferentes CBDC de cada país para así hacer importaciones pagando con la CBDC de tu país.

Antes de explicar los tipos de CBDC hace falta hacer otra aclaración para evitar confusiones que puedan aparecer. El *token money* o dinero simbólico (propiedad tangible) es un tipo de dinero que por definición tiene un valor nominal superior a su valor intrínseco, y a diferencia del dinero en cuenta (propiedad intangible), no es un número presente en la contabilidad de un banco sino que es el dinero que la población usa. Ejemplo de esto lo encontramos en los billetes y monedas de curso legal que hay en circulación hoy en día (Euro, Dólar, etc.). El euro en 1999 se convirtió en dinero en cuenta y no fue hasta 2002, cuando empezaron a emitirse monedas y billetes, que se convirtió en dinero simbólico ("Token Money", s.f., párrafo 1).

Las CBDC podrían ser de dos tipos, basadas en cuentas o en tokens (Griffoli et al. 2018):

- **Basado en cuenta:** sería semejante a la situación que hay hoy en día, en la que puedes llevar a cabo una transferencia entre dos entidades bancarias, con la única diferencia de que las cuentas estarían en el banco central. Si alguien quiere llevar a cabo una transacción

se metería en su cuenta del banco central y entonces pediría una transferencia de fondos a otra cuenta receptora. El banco central tendría que encargarse de cuatro aspectos:

- Verificar la autoridad por parte del remitente para llevar a cabo la transacción.
 - Verificar que posee suficientes fondos.
 - Autenticar la autenticidad del receptor de fondos.
 - Una vez se ha validado la transacción actualizar el libro mayor que contiene todas las transacciones.
- **Basada en tokens:** cuando se dice basada en tokens, no se refiere a los tokens explicados anteriormente. Con esto se hace referencia a que las CBDC serían instrumentos para los que no es necesario registrar la propiedad ante ninguna persona jurídica o entidad, ya que la propiedad se denota por la capacidad que tiene una persona de realizar alguna acción, como por ejemplo producir una firma digital. Los bancos centrales podrían permitir la existencia de llaves públicas y privadas sin la necesidad de que haya una identidad asociada a cada una de ellas, pero limitando su uso de acuerdo a la legalidad vigente de cada país. Por lo tanto, a diferencia del dinero en efectivo de hoy en día, los CBDC tokens tienen una serie de diferencias:
 - Son más difíciles de distinguir de falsificaciones por parte de las personas que llevan a cabo la transacción.
 - Las transacciones, por lo tanto, requerirían una verificación externa.
 - Las transacciones no serían enteramente anónimas como pasa con el dinero en efectivo. El nivel de anonimidad, tal y como aclara en (Griffoli et al. 2018), dependería de si las billeteras están registradas o si la información de las transacciones está grabada.
 - La verificación se podría llevar a cabo mediante:
 - * Un libro mayor distribuido (por ejemplo mediante una blockchain no permissionada), pero la contrapartida es que esta tecnología tiene problemas de escalabilidad (aplicarse a gran escala y ser capaz de procesar millones de transacciones) (He et al. 2017 citado en Danielsson (2019)). Tal y como se afirma en (Danielsson 2019), si no hay una mejora significativa en la tecnología Blockchain que permita el procesamiento a gran escala de millones de transacciones, esta opción es inviable.
 - * Una red permissionada manejada por el banco central (parecido a la blockchain permissionado de Facebook para Libra).

En (Griffoli et al. 2018) se da información de varias encuestas que se han hecho a bancos centrales y estas son las principales conclusiones:

- Algunos mencionan que en caso de existir las CBDC deberían estar disponibles las 24 horas el día todos los días de la semana para así imitar la accesibilidad del dinero.

- Ninguno de los bancos centrales están considerando seriamente que las CBDC tengan intereses. Estos están preocupados por la intermediación financiera y la mayor volatilidad del balance bancario. Los bancos centrales que tienen el foco en un CBDC basado en tokens afirman que pagar intereses podría suponer un desafío técnico difícil de superar.
- La mayoría de los bancos centrales están centrados en una aplicación doméstica de las CBDC. Canadá está considerando que sean accesibles por turistas y China menciona problemas relacionados con la gestión de capital transfronterizo. Finalmente cabe destacar que algunos proyectos se centran más en la intermediación de pagos mayoristas a través de las fronteras, como un proyecto conjunto entre el Banco de Canadá, la Autoridad Monetaria de Singapur y el Banco de Inglaterra.

Una vez se ha entendido cual sería el funcionamiento de las CBDC hace falta destacar dos puntos negativos que pueden hacer inviable que estas lleguen a existir (Danielsson 2019):

- Los ciudadanos, al estar haciendo las transacciones en la blockchain del banco central, éste se convierte en un proveedor de servicios y por lo tanto ha de establecer toda la infraestructura necesaria para ello (centros de llamadas, cableado, equipos informáticos, software, etc.).
- El banco central, como pasa a ser el receptor del dinero de la población, sustituyendo la función de los bancos comerciales, entonces se convierte en el responsable de dar créditos a los mismos. En (Danielsson 2019) afirma que esto es algo que un banco central no querría tener que realizar ni es algo que vayan a hacer, ya que no forma parte de sus funciones.

4 Coexistencia del dinero fiat y las criptomonedas

En la historia ha habido casos como el del bimetalismo, sistema monetario que ha tenido más presencia alrededor del mundo, el cual ha tenido sus diferencias según la época y país, pero parte de la premisa de que el valor de una unidad monetaria está definida con una equivalencia respecto a dos metales, principalmente el oro y la plata, teniendo estos dos últimos un tipo de cambio fijo entre ellos (“Bimetallism”, s.f.). El problema al que se ha enfrentado este sistema repetidamente es la Ley de Gresham, la cual afirma que (“Ley de Gresham”, s.f.):

Cuando en un país circulan simultáneamente dos tipos de monedas de curso legal, y una de ellas es considerada por el público como “buena” y la otra como “mala”, la moneda mala siempre expulsa del mercado a la buena. En definitiva, cuando es obligatorio aceptar la moneda por su valor facial, y el tipo de cambio se establece por ley, los consumidores prefieren ahorrar la buena y no utilizarla como medio de pago.

Bitcoin, la criptomoneda más usada a nivel mundial, tiene una serie de problemas que pueden hacer que esta consideración de coexistencia quede totalmente descartada no por la incompatibilidad entre ellas, o por la Ley de Gresham, sino por el poco uso transaccional fuera de lo legal que se le da a Bitcoin, que puede ser aplicado al resto de criptomonedas que existen en la actualidad. La población las compra para especular con ellas, y esto, junto a la volatilidad de precios, hace que el uso transaccional que se les pueda dar desaparezca (Zimmerman 2020 citado en Danielsson (2019)).

Para entender la argumentación que da (Zimmerman 2020), antes hay que entender como de limitada está la oferta de Bitcoins, la criptomoneda más utilizada. En la figura 6 se puede ver la evolución del número de bitcoins que existen en millones de 2009 (año de su creación) a 2023. Cada cuatro años el problema computacional que los miners resuelven da como recompensa la mitad de Bitcoins (empezando en 50 bitcoins en 2009) y se vuelve cada vez más complicado. El límite teórico, el cual se alcanzará en 2140 (Danielsson 2019), será de 21 millones de monedas. Con tal de verificar las transacciones y limitar el número de monedas existentes, los bitcoins solo pueden ser creados mediante la actividad del mining, que como se ha comentado en apartados anteriores, compensa a aquel que resuelve el problema con un número x de Bitcoins, los cuales son de nueva creación. Esta es la **única** manera de crear Bitcoins. Tal y como se comenta en (Danielsson 2019), el hecho de que cada vez sea más complicado resolver el problema hace que manipular la blockchain sea más improbable a medida que pasa el tiempo, aspecto que hace que aumente la confianza.

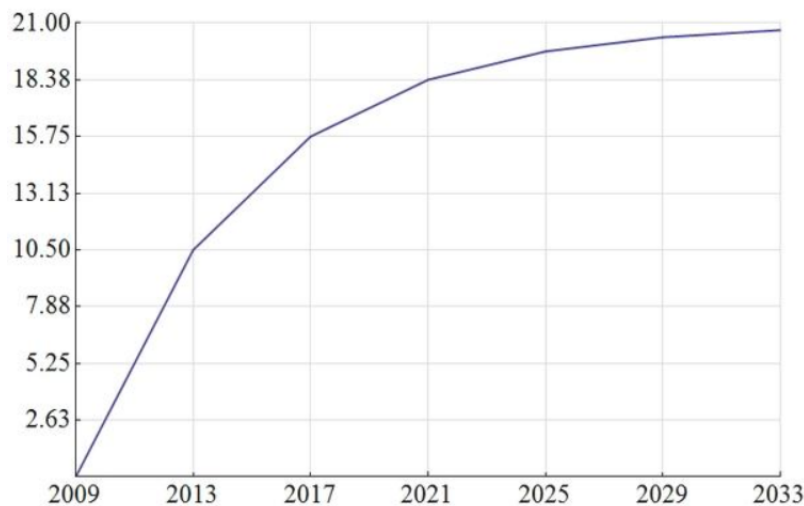


Fig. 6: Millones de bitcoins existentes desde su creación junto a la progresión que tendrán hasta 2033. Fuente: ¿Son 21 millones de bitcoins suficientes?

Aunque el resto de criptomonedas (Ethereum, Ripple, Monero, etc.) tienen una estructura parecida de creación de monedas, hay algunas diferencias que suelen presentarse (Danielsson 2019):

- Algunas crean cierta cantidad de la nada mediante un proceso que se llama *pre-mining*, que se da en lo que se llama oferta inicial de monedas. Tal y como se define en ("Oferta inicial de monedas", s.f.), "las criptomonedas son vendidas en forma de "tokens" a especuladores o inversores a cambio de dinero tradicional u otras criptomonedas como Bitcoin o Ethereum. Los tokens son vendidos como "futuras" unidades de la moneda cuando la ICO llegue a su objetivo y el proyecto se lance".
- Hay otras criptomonedas que tienen un proceso de creación más rápido que el de Bitcoin, lo que permite que haya transacciones más rápidas debido a que el problema computacional a resolver, el cual permite registrar las transacciones en una blockchain y por lo tanto

confirmarlas, es más rápido.

Una vez se ha entendido esto se procede a la explicación de (Zimmerman 2020) de por qué tienen poco potencial como medio transaccional:

- Las criptomonedas tienen una característica distintiva respecto a otros activos y esta es la endogeneidad del precio, el cual viene determinado por el uso que se le da a la moneda y depende del valor que la sociedad considera que tiene, ya que estas no tienen valor intrínseco.
- Un pago mediante criptomonedas no queda confirmado hasta que no está registrado en una blockchain. Como el espacio es limitado, hay mucha competencia entre poseedores de las mismas para confirmar sus transacciones, la cual puede verse incrementada de una forma muy importante en momentos de pánico. Si hay una gran presión especulativa, las tarifas asociadas a cada transacción, las cuales establece el remitente para dar más o menos prioridad a su transacción (los mineros, como son compensados con parte de estas tarifas, incluyen las más rentables en el bloque para el que van a resolver el problema computacional), se ven incrementados mucho en precio, ya que el número de transacciones que se pueden incluir en un bloque es limitado. Es por hechos como este que la habilidad de una criptomoneda para servir como medio de pago se ve perjudicada y por lo tanto reduce su valor y por ende su precio (Zimmerman 2020).

Es por razones como esta que las criptomonedas van a tenerlo muy complicado para estar presentes en nuestra actividad del día a día. El dinero fiat está presente en los contratos laborales, en los préstamos y en todos los aspectos de la economía. En (Zimmerman 2020) hay una frase que define muy bien esta situación: “I don’t think many people would like to earn their salaries in dollars, pay rent in Bitcoin, buy groceries with Ethereum, and compensate the hairdresser in Ripple”. La población generalmente lo que quiere es usar una única moneda, que sea estable en precio, que su precio sea más o menos predecible y que permita realizar transacciones con facilidad. Esto está muy lejos de darse en las criptomonedas y por lo tanto deja como una opción muy remota la posible coexistencia.

Pese a todo esto, lo que sí que es una realidad es que puede darse la situación de que su uso se vea muy incrementado y sean de utilidad en países pocos desarrollados, con población sin acceso al sistema bancario, con unas monedas poco estables y manejadas de una manera muy impredecible por el banco central, donde un sistema basado en criptomonedas podría tener ventajas sobre uno basado en dinero fiat, sobre todo para llevar a cabo transferencias de dinero. Esto es algo que Facebook ha sabido ver y de ahí viene principalmente la creación de Libra como un medio para ayudar a los países menos desarrollados.

5 ¿En qué sistema monetario estaría basado un criptosistema?

En caso de que llegase a existir un criptosistema cabe preguntarse que sistema monetario tendría. Para explorar esto primero hay que definir que sistemas monetarios existen, los cuales son de reserva fraccionaria o de reserva completa (Danielsson 2019):

- En un sistema de reserva fraccionaria, el total de dinero que hay en manos del público o

base monetaria (BM), el cual es el dinero creado por el banco central, no es igual al total de dinero en circulación, y esto es debido a que por ejemplo, en la Eurozona, la banca solo tiene que mantener un porcentaje del dinero en reserva cercano al 1% en el caso de Europa. Esto significa que cuando tú depositas 1000 euros en tu cuenta bancaria, el banco tienen un requerimiento de tener un x% guardado (requerimiento de reserva), supongamos que es un 10% o equivalentemente 100 euros. Los 900 euros restantes el banco los puede usar para prestarlos a una persona, con lo que en total habría 1900 euros disponibles en la economía, de los cuales solo 900 existen en físico. Si este proceso continua, y si un banco parte de solo 1000 euros y con un requerimiento de reserva del 10%, la oferta de dinero total (M1) que habrá será de $\frac{Cantidad\ inicial}{Requerimiento\ de\ reserva} = \frac{1000}{0.1} = 10000$ euros. El dinero, por lo tanto, tiene diferentes medidas que se definen a continuación ("Masa monetaria y dinero ¿qué son M0, M1, M2, M3 y M4?" ,s.f.,párrafos 6-14):

- **M0:** este es el dinero que circula en la economía o base monetaria, que es igual a los billetes y monedas en manos de los ciudadanos, además del dinero que los bancos tienen en sus cajas y depositado en el banco central. Dicho de otra forma, es el dinero en efectivo.
 - **M1 u oferta monetaria:** es el dinero que circula en la economía. Esto incluye M0 y los depósitos corrientes de los ciudadanos y empresas.
 - **M2:** incluye M1 y los depósitos de ahorro.
 - **M3:** incluye M2 y los depósitos bancarios a más largo plazo y los depósitos en moneda extranjera.
 - **M4:** incluye M3 y las emisiones de bonos, letras del tesoro y pagarés.
- En un sistema de reserva completa los bancos mantienen el 100% del dinero de los depósitos en efectivo y por lo tanto no pueden llevar a cabo préstamos con el dinero depositado. Tal y como se comenta en (Danielsson 2019), es un sistema que da más estabilidad y seguridad, no se pagan intereses por tener el dinero en el banco pero no se evita la creación de M1, M2 y M3, ya que la población tendría la opción de depositar el dinero en bancos con menos regulación o instituciones con menos transparencia.

En un criptosistema basado en un sistema de reserva completa los dueños de criptomonedas querrían prestar a otros a cambio de alguna retribución. Siempre y cuando esto pase entre dos individuos o a través de instituciones financieras que operen bajo reserva completa, esto es posible. El problema está en que las instituciones financieras a través de las cuales se realizan las transacciones no operen con reserva completa, es en esta situación que se crearía cripto M1, M2, M3 y M4. Esto sería prácticamente inevitable y por lo tanto las transacciones dejarían de estar controladas en una blockchain. Por lo tanto, se acabarían teniendo bancos con reservas de cripto, instrumentos de crédito cripto, derivados cripto, tal y como sucede con el dinero fiat. Por lo tanto, un sistema de reserva completa es muy improbable.

El problema que esto supone es que habría una gran cantidad de M1, M2 y M3 cripto que serían reclamaciones de criptos y no monedas. Por lo tanto, ese aspecto en el que destacan las criptomonedas, que es en el uso de la tecnología blockchain para registrar todas las transacciones, dejaría de ser algo que refleje la realidad de las transacciones que se realizan

con las criptomonedas y por lo tanto perderían ese hecho distintivo y se convertirían en una alternativa muy parecida al dinero *fiat* de curso legal. Además, esto supone una serie de problemas financieros que serán explicados más adelante.

6 La estabilidad de precios

Con tal de saber como de bien funciona un sistema monetario hace falta saber como es capaz de lidiar con los principales problemas adversos que puede enfrentar, como son: la inflación, la deflación y el despalancamiento. Este último se da en periodos de crisis cuando los hogares, empresas e inversores que se habían endeudado con anterioridad no son capaces de soportar las deudas que tienen, ven como su solvencia se reduce y esto hace que lleven a cabo un desendeudamiento rápido en unas condiciones que son desfavorables (“Desapalancamiento”, s.f.). Un ejemplo de esto sería vender una casa que tenías financiada con un préstamo.

La inflación a niveles inferiores al 2% se considera una inflación aceptable, ya que incentiva el consumo (la población no guarda los ahorros con la expectativa de que los precios van a bajar como pasa con la deflación). Ésta se ha descontrolado en numerosos momentos de la historia, afectando consecuentemente a las monedas *fiat*. Algunos ejemplos de esto son la hiperinflación en Alemania debido al Tratado de Versalles, que ante la incapacidad de pagar las reparaciones, entre otras cosas, hizo que en 1923 llegara a tener una inflación del 1 000 000 000 000%. Otro caso más cercano en el tiempo es el de Venezuela, que de 2013 a 2019 ha tenido una inflación acumulada del 65 395 563 286% (“Hiperinflación”, s.f.).

Tal y como se menciona en (Danielsson 2019), muchas personas a favor de las criptomonedas argumentan que una oferta fija de criptomonedas garantiza que haya estabilidad de precios. Hay una teoría, que es la teoría cuantitativa del dinero, que pone esto en cuestión. Está definida de la siguiente manera (“Teoría cuantitativa del dinero”, s.f.):

$$\text{Nivel de precios} = \frac{\text{cantidad de dinero} \cdot \text{velocidad de circulación}}{\text{nivel de producción}}$$

En consecuencia, el nivel de precios aumenta en estas situaciones:

- Cuando la oferta de dinero se incrementa y la velocidad de circulación y el nivel de producción se mantienen constantes.
- Cuando la velocidad de circulación se incrementa y el nivel de producción y la cantidad de dinero se mantienen constantes.
- Cuando hay una reducción de la actividad económica y la oferta de dinero y la velocidad se mantienen constantes.

Cabe decir que han surgido teorías alternativas como la de Milton Friedman que afirma que a largo plazo las variaciones de la oferta de dinero no tienen ningún efecto en la economía y que el único efecto que se produce es a corto plazo y es sobre el nivel de precios.

El problema existente que hace que ni un sistema monetario basado en dinero *fiat*, en oro o en criptomonedas garantice la estabilidad de precios es que estos componentes no son determinados por si solos si no que son determinados conjuntamente en el equilibrio. Por lo tanto, ante un

cambio en la oferta de dinero, los efectos en los precios son indeterminados porque tanto la velocidad del dinero y la cantidad de dinero cambian a causa de ello. Además, como esta oferta de dinero es $M1 + M2 + M3 + M4$, y estos tipos de dinero dependen a su vez de $M0$, que es el que emite el banco central, controlar perfectamente los cambios que se van a producir en los tipos superiores de dinero ($M1$ hacia arriba) es muy difícil.

Poniendo el foco en un sistema monetario basado en criptomonedas, concretamente en Bitcoin, debido a que ésta tiene un aumento lento de su número (se crean cada 10 minutos en promedio, 50 bitcoins en 2009, y a partir de allí, la mitad cada cuatro años), lo más seguro es que un criptosistema basa en esta criptomoneda sufra de deflación (Danielsson 2019). Esto es debido a que si la actividad de minería es más lenta que el crecimiento económico, a no ser que la velocidad se incremente continuamente, que no es el caso, los precios bajarán.

El caso más semejante a esto se puede ver en el patrón oro que hubo entre 1872 y 1914. Debido al hecho de que la oferta de dinero no creció al ritmo de la economía, los precios cayeron y por lo tanto esto benefició a los que eran propietarios de capital (los ahorradores veían como su dinero ganaba valor a medida que pasaba el tiempo) y perjudicó a los trabajadores y los prestatarios (los salarios nominales bajaban continuamente y los préstamos se encarecían) (Danielsson 2019).

En consecuencia se ha de garantizar que si la economía está creciendo a un ritmo elevado, la oferta de dinero tiene que hacerlo en concordancia con tal de prevenir la deflación. Muchas recesiones, tal y como se menciona en (Danielsson 2019), vienen acompañadas de desentplancamiento, y esto se puede ver en incrementos de las formas de dinero $M0$ y $M1$ y un incremento de $M2$ y $M3$. Es por ello que incrementar la oferta de dinero durante una recesión previene daños mayores.

En la figura 7 se ve como éstas evolucionaron durante la Gran Depresión. Como se puede observar, en el periodo que la comprende, el cual está marcado en amarillo, hay una reducción constante de $M1$ y $M2$ y por lo tanto estaba habiendo un desentplancamiento a la vez que $M0$ se incrementaba a un ritmo lento. De hecho, esta es la explicación que dio Friedman para explicar la Gran Depresión desde un punto de vista monetarista.

John Maynard Keynes dio una explicación distinta, asociando la misma a una bajada de la curva de eficiencia marginal del capital (beneficios esperados de los empresarios no frutos de un cálculo, ya que, bajo su visión, la población lleva a cabo las decisiones bajo condiciones de incerteza). Esta bajada provocó una disminución de la inversión y consecuentemente del PIB. Esto a su vez provocó un aumento de la preferencia por la liquidez (aumento de la demanda de dinero), que hizo incrementar el tipo de interés y consecuentemente la inversión bajó aún más. Él denominó Animal Spirits al hecho de que los individuos toman las decisiones según el optimismo espontáneo que estos tienen y no según las expectativas racionales (Friedman contrariamente a esto pensaba que la población era muy racional).

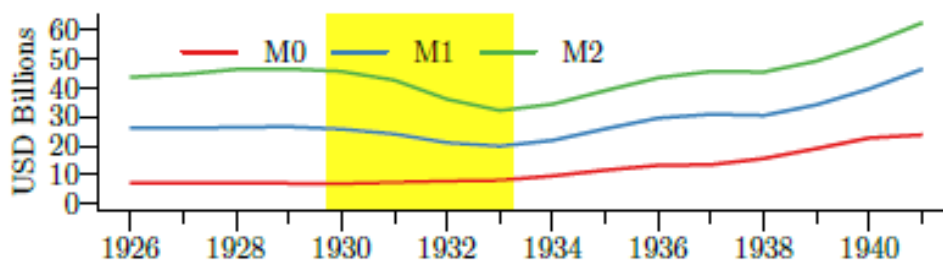


Fig. 7: Evolución de M0, M1 y M2 durante la Gran Depresión, 1929-1933. Fuente: (Danielsson 2019)

Sea o no ésta la causa principal, lo que está claro es que es que este desenganpalamiento que se produjo, el cual según Friedman se tendría que haber tratado de solventar mediante aumentos de la oferta monetaria por parte de la Reserva Federal, ralentizó de una forma muy importante la economía.

Por lo tanto, la conclusión que se puede sacar de este apartado es que ninguno de los tres sistemas monetarios mencionados son una garantía de estabilidad de precios, ya que todos ellos pueden generar tanto inflación como deflación. La gran diferencia está en el hecho de que con un sistema monetario basado en dinero fiat el banco central tiene la capacidad para ajustar la oferta de dinero según la situación económica en la que ese esté, mientras que con un criptosistema, ésta posibilidad no existe, ya que, en este último, la creación de monedas suele estar basada en un algoritmo que las crea a tasas constantes o decrecientes y bajas. En consecuencia, la estabilidad de precios, tal y como apunta en (Danielsson 2019), está más garantizada con un sistema basado en dinero fiat bien manejado, ya que en caso contrario, como en Venezuela, puede haber hiperinflación.

7 La estabilidad financiera de un criptosistema

Aunque a primera vista se puede llegar a pensar que las criptomonedas no suponen una amenaza a la estabilidad financiera, ya que su peso en las transacciones que se dan hoy en día es muy pequeño en comparación con el uso que se le da al dinero fiat, además de que se utilizan primordialmente con fines especulativos y no tienen ningún uso económico de relevancia, esta situación cambia en caso de que se diera un uso a nivel mundial de las mismas. Podrían pasar a ser uno de los actores principales en la economía y es entonces cuando podrían aparecer problemas financieros de estabilidad que todos los sistemas financieros tienen. Tanto los sistemas basados en dinero fiat como en oro siempre han tenido el problema de que pese a que el banco central puede tener bajo control M0, en tiempos de crisis la creación de M1, M2 y M3 no se puede controlar con facilidad (Danielsson 2019). Cuando hay prosperidad económica, estas tres últimas formas de dinero suelen aumentar debido a que se dan más créditos y cuando hay recesión estas altas formas de dinero son convertidas en formas más bajas, provocando una crisis.

En (Danielsson 2019) se mencionan varios ejemplos en los que esto ha ocurrido. Durante el patrón

oro en Inglaterra, a mediados del siglo 19, el banco de Inglaterra tenía suficiente oro para cubrir el 40% del dinero impreso. Este dinero impreso fue convertido en formas más altas de dinero como lo son M1 y M2 y cuando sufrió una crisis económica en 1866, la población se apresuró en convertir sus activos financieros en el activo más líquido, en este caso oro. Esto supuso una caída muy importante de la oferta de dinero, una crisis crediticia y bancarrotas en todo el país. Es por crisis como ésta que apareció el prestamista de última instancia, proveído por los bancos centrales, cuya labor consistía en garantizar liquidez de los bancos en tiempos de retiradas masivas o imprevistas, para de esta forma hacer desaparecer el miedo a no poder retirar el dinero en tiempos de crisis ("Prestamista de última instancia", s.f.).

De igual forma, el autor señala que la misma fuerza fundamental que causó la crisis de 1866 en Inglaterra está detrás de casi todas las crisis financieras, incluida la de 2008, y esta es la cantidad excesiva de riesgo endógeno. Este concepto fue propuesto por Jon Danielsson y Hyun Song Shin en 2002 ("El riesgo endógeno - Endogenous risk", s.f.).

El riesgo puede ser clasificado en dos categorías, exógeno (perturbaciones que llegan desde fuera del sistema) o endógeno, debido a la interacción de los participantes en el mercado, los cuales tienen sus propias habilidad, prejuicios y recursos. Este riesgo endógeno a la vez puede ser subdividido en riesgo real, que es el latente, y el riesgo percibido, que es el que se informa mediante las técnicas de medición de riesgo más comunes.

En la figura 8 se puede ver una representación de los mismos en un caso genérico. Durante el periodo de burbuja (1-12), donde los precios no paran de subir, el riesgo percibido disminuye y el riesgo real se va incrementando. En este periodo de boom la población se comporta de forma más idiosincrática respecto al periodo de crisis, que sucede cuando la burbuja estalla, causando que la población actúe al unísono guiados por un pánico generalizado que crece sin control alguno. Todo aquel que cree estar en riesgo real, ya que percibe un riesgo elevado tiene el instinto de preservar lo que uno tiene. Los precios bajan de una forma muy brusca al mismo tiempo que el riesgo real disminuye y el riesgo percibido aumenta. La conclusión que se saca por lo tanto es que la percepción del riesgo y el riesgo real tienen correlación negativa.

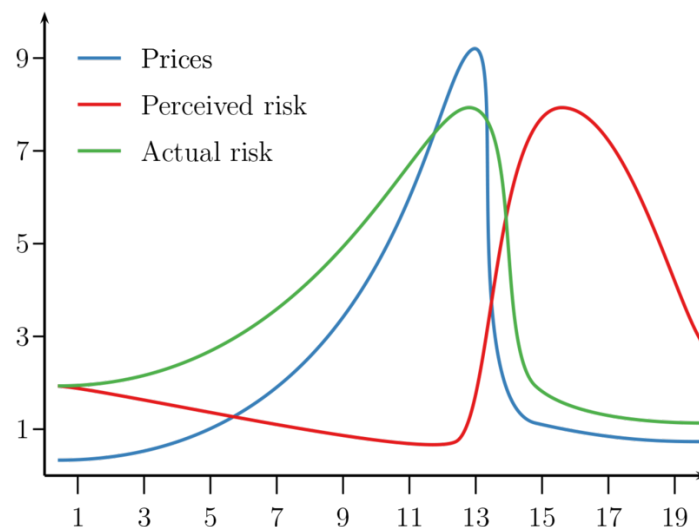


Fig. 8: Riesgo real y riesgo percibido. Fuente: El riesgo endógeno - Endogenous risk

Tal y como se sugiere en (Danielsson 2019), un criptosistema es más proclive a este riesgo endógeno que un sistema monetario basado en dinero fiat u oro. Esto es debido a que la estabilidad que un hipotético criptosistema tendría (debido a que, en caso de que la criptomoneda usada fuera Bitcoin, el incremento de la oferta estaría definido y sería inquebrantable), provocaría más desestabilización que si fuera más inestable. Esto es debido a que cuanto más estabilidad transmita un sistema monetario más riesgo van a correr los agentes económicos y por lo tanto esto implica mayores inversiones en proyectos y endeudamientos de mayor magnitud y por lo tanto ante un pequeño evento puede haber consecuencias mucho más graves comparado con la situación en que la población hubiera tomado más precauciones.

Otro de los problemas, esta vez fuera de interpretaciones sobre cual es la estabilidad de la que se presume que tendría un criptosistema, y más perceptible, es el asociado a los cripto M1, M2 y M3. Al igual que en un sistema monetario basado en dinero fiat, se crearían estas formas más elevadas de dinero, la gran diferencia está en que contrariamente a lo que pasaba con el patrón oro que hubo entre 1944 y 1973, no habría ninguna garantía de convertibilidad en oro o en otras monedas fiat. Esta cripto M1, M2 y M3 lo más seguro es que serían simplemente reclamos de monedas (Danielsson 2019), esto haría que el pánico fuera mucho más importante. ya que la población empezaría a preocuparse sobre si el dinero que tiene es equivalente a las criptomonedas en forma de M0 que existen.

Por lo tanto, a no ser que sea un criptosistema basado en una criptomoneda distinta a Bitcoin, donde la “encuñación” de las monedas no se haga a una tasa que no puede ser flexible, este problema puede hacer inviable la existencia de un criptosistema. Si algo se ha aprendido de la historia es que en momentos de pánico el banco central ha de garantizar la liquidez de los bancos para que, por un lado, en momentos de crisis no haya un pánico generalizado o sencillamente para estimular la economía. Si un criptosistema no dispone de esta herramienta, entonces se enfrenta a la cruda realidad de que ante una crisis económica no pueda responder de la manera adecuada para garantizar la estabilidad.

8 Las transferencias de dinero

Uno de los problemas que enfrenta Bitcoin es el coste asociado a las transacciones. Éste, a lo largo de los últimos años, ha estado entre los 0.45 dólares y los 55 dólares (Danielsson 2019). Como ya se ha comentado con anterioridad, esto sirve como incentivo para que aquellas empresas o particulares que destinan mucho dinero en equipos informáticos para realizar la actividad del *mining* y así poder ser compensados por ello, tengan un incentivo para continuar expandiendo la blockchain y garantizar así la seguridad que tiene asociada. El problema que existe es que cuando se lleva una transacción, esta puede llegar a ser millones de veces más pequeña que un Bitcoin (este tiene un precio actual de 6262,26 euros a 12 de abril de 2020) y por lo tanto, si uno no quiere realizar una transacción de grandes cantidades, ha de soportar un coste para tener su transacción añadida a la blockchain que puede ser demasiado elevado como para que compense y que además puede llevar de 10 minutos a una hora en llevarse a cabo.

Es por situaciones de pánico como ésta que hacen falta otras alternativas que garanticen un procesamiento rápido de las transacciones, aunque esto, por motivos obvios, haría que la

seguridad se viera perjudicada si el método de mining es el mismo que el de Bitcoin. Hasta ahora solo hay propuestas sobre esto, pero ninguna se ha realizado (Danielsson 2019).

Si esto se compara con una transacción usando una sola moneda fiat, uno se da cuenta de las deficiencias que las criptomonedas tienen. Hoy en día, por ejemplo, en Europa, puedes enviar dinero de manera barata, segura y prácticamente instantánea a cualquier país. Esto sucede en muchos otros países gracias al e-money (PayPal, M-Pesa, etc.). Pese a todo esto, el dinero fiat presenta problemas de altos costes y procesamiento de transacción lento cuando se trata de transacciones que implican el intercambio de monedas. Esto es debido a que hay una serie de leyes como las leyes antiblanqueo o lo conocido como “conozca a su cliente”, que es el proceso de una empresa para identificar y verificar la identidad de los clientes (“Conozca a su cliente”, s.f.).

El único ámbito en el que las criptomonedas o las monedas digitales corporativas tienen potencial para ser útiles es en las transacciones que suceden en países en desarrollo, ya que es en estos donde hay por un lado, mayores costes asociados a las mismas y además hay millones de personas sin posibilidad de acceder al sistema bancario, tal y como se menciona en la página de Libra, moneda digital que quiere crear Facebook.

9 ¿Por qué Libra no va a ser lo que pretendía?

Aunque Libra parecía un proyecto con mucho potencial, este está viendo como una serie de obstáculos que están posponiendo su puesta en marcha o incluso imposibilitándola. Ueli Maurer, ministro de finanzas de Suiza, afirmó en diciembre de 2019 a la emisora Suiza SRF que no piensa que Libra en el estado actual tenga una oportunidad de establecerse como una moneda digital corporativa debido a que los bancos centrales no van a aceptar que haya una cesta de monedas fiat que la sustente (“Facebook’s Libra cryptocurrency project has failed in its current form, says Swiss president”, 27 de diciembre de 2019, párrafos 1-2).

Debido al hecho de que Libra sería soportada en todas las apps de Facebook, incluyendo Instagram y Whatsapp (suponiendo más de 2000 millones de usuarios), esto haría que fuese una competencia muy importante con monedas establecidas como el dólar estadounidense o el euro y en consecuencia podría causar una desestabilización del sistema financiero actual. Es por motivos como este que también ha tenido la oposición del ministro francés de economía y finanzas, Bruno Le Maire, quien dijo que suponía una amenaza a la soberanía monetaria. Por otro lado, en julio de 2019, políticos de Reino Unido y los Estados Unidos expresaron preocupación, llegando a afirmar durante el comité de banca, vivienda y asuntos urbanos del senado de los Estados Unidos que Facebook no merece su confianza debido a los escándalos de privacidad de los recientes años (“Facebook’s Libra cryptocurrency project has failed in its current form, says Swiss president”, 27 de diciembre de 2019, párrafos 8-9).

Esta tiene una serie de implicaciones sobre la estabilidad financiera que han sido algunas de las responsables de su rechazo (“Libra, la criptomoneda de Facebook”, 4 de octubre de 2019, párrafos 13-17):

- La estabilidad de Libra depende de que los activos que la respaldan tengan suficiente estabilidad.

- Podría contribuir a generar episodios globales de inestabilidad financiera. En caso de que hubiera una retirada masiva de libras por un pánico generalizado, debido al hecho de que la Asociación pretende invertir las divisas que obtiene de la venta de libras en activos de bajo riesgo, cabría estudiar si sería capaz de satisfacer la demanda en caso de que estos activos no tuvieran la suficiente liquidez.
- Aunque Libra puede tener algunos efectos positivos como por ejemplo facilitar los flujos de capitales internacionales debido a la reducción de los costes asociados a las transacciones que habría, puede afectar la estabilidad financiera de economías emergentes, podría empeorar las salidas de capitales de las mismas en caso de pánico.
- En países con monedas poco estables, donde Libra puede tener mayor efecto, podría condicionar la política monetaria de sus bancos centrales. En caso de que los residentes tengan la capacidad de cambiar su moneda local por una cesta estable de divisas a través de Libra, en caso de la mínima estabilidad económica en su país se refugiarían en éstas, por lo que la moneda local podría sufrir grandes depreciaciones (aumentaría la oferta de la moneda local), afectado de esta manera la estabilidad de precios.

Desde que en junio de 2019 Facebook anunció los planes para el desarrollo de Libra, ésta estaba asociada con proveedores de pagos como Visa, Mastercard y PayPal. A los pocos meses de este anuncio estas tres empresas abandonaron el proyecto junto a otras muchas. Las únicas empresas tecnológicas de relevancia que continúan en la Asociación Libra son Uber, Lyft y Spotify debido a que estas venden en Facebook (“Why Facebook’s Libra Is not a Cryptocurrency, and Why It Is Doomed to Fail”, 28 de enero de 2020, párrafos 1-3).

Esta moneda digital ha generado preocupación entre los reguladores y políticos debido a su privacidad y el potencial que tiene para influir en la política monetaria y los cambios que esto podría suponer en el panorama financiero global. De hecho, David Marcus, cocreador de Facebook, afirmó que un conjunto de obstáculos reglamentarios puede hacer que la fecha de lanzamiento se posponga más allá de junio de 2020 (“Facebook’s Libra cryptocurrency project has failed in its current form, says Swiss president”, 27 de diciembre de 2019, párrafo 4“)

Pero el cambio importante es que, debido a las enormes presiones y oposición que ha tenido, ha cambiado totalmente su funcionamiento (“Facebook is shifting its Libra cryptocurrency plans after intense regulatory pressure”, citando a “Facebook Revamps Libra Plans, Bowing to Regulators”, 3 de marzo de 2020, párrafos 1-2). Facebook ya no tiene la intención de que Libra sea el centro de la estrategia de pago. En vez de esto, el proyecto soportará transacciones con otras monedas fiat como son el dólar estadounidense y el euro además de con la moneda digital corporativa Libra.

Ahora, Libra tiene la intención de tener un conjunto de monedas estables digitales que denomina “single-currency stablecoins” (SCS) que representan diferentes monedas nacionales como son el euro o el dólar, pero en una forma digital simbólica o tokenizada. Además, pretende que a medida que pase el tiempo cada vez haya más monedas estables digitales que representen monedas nacionales de otros países (“Economics and the Libra Reserve”, s.f., párrafo 3). Esto permitirá que los negocios y particulares lleven a cabo transacciones con una SCS denominada en su propia moneda.

En vez de tener una reserva de efectivo y valores gubernamentales más seguros a corto plazo que soportarán el valor de Libra, ahora existirá esto mismo pero para Libra y todas las SCS. La diferencia está en que ahora cada SCS tendrá valores gubernamentales del país de origen de la moneda que representan, y lo mismo pasa con el efectivo que haya en la reserva. Las SCS se acuñarán y se quemarán en respuesta a la demanda del mercado para esa moneda y debido al respaldo 1: 1 de cada moneda, no habrá una creación de dinero neto (“Economics and the Libra Reserve”, s.f., párrafo 3).

Esto se ha hecho con el objetivo de dar más flexibilidad y control a los bancos centrales para controlar el uso del sistema de pagos Libra en cada uno de sus países. Para aquellos países que no dispongan de una SCS en la red Libra, tienen la opción de utilizar Libra debido a su poca volatilidad y al hecho de que puede aumentar la inclusión financiera de sus habitantes. Podría funcionar como una alternativa para las transacciones transfronterizas, pudiendo, las personas y negocios que reciben Libras, convertirlas en su moneda local.

En (“Economics and the Libra Reserve”, s.f., párrafo 6) deja claro cual sería el funcionamiento de la red Libra:

- Si una persona que vive en los Estados Unidos quiere enviar US\$, y el receptor vive en un país que tiene una SCS distinta, el remitente tiene dos opciones:
 - puede enviar el dinero con la SCS del país del receptor.
 - el remitente puede enviarlo con la SCS del país en el que está, en este caso en SCS que representan los US\$, y el receptor convertirlo en su SCS o en su moneda local a través de un proveedor de servicios financieros externo.
- Si una persona que vive en los Estados Unidos quiere enviar US\$ pero en el país receptor no hay ninguna SCS, la transferencia se puede realizar en monedas Libra, y el receptor puede convertir la Libras en su moneda local a través de un proveedor de servicio financiero externo.

Por lo tanto, Libra no provee ni establece ningún tipo de conversión entre las monedas Libra y el resto de monedas fiat sino que es a través de proveedores de servicios externos que se lleva a cabo esta conversión.

Como Libra no está vinculada a una sola moneda, sino que es una “moneda multdivisa” consistente en un compuesto digital de algunas de las SCS de la red Libra, donde cada SCS tendrá un peso determinado en la determinación del valor de Libra. Cuando el valor de estas monedas fluctúe, el valor de las monedas Libra en cualquier moneda local también fluctuará. Tal y como se menciona en (“Economics and the Libra Reserve”, s.f., párrafo 15):

La Asociación agradecería la supervisión y el control sobre la moneda Libra por parte de un grupo de reguladores y bancos centrales o una organización internacional (por ejemplo, el FMI) bajo la guía de la principal autoridad supervisora de la Asociación, FINMA, que podría supervisar y controlar los pesos y componentes para minimizar la volatilidad.

Independientemente de la región, esperan exigir a todos los proveedores de servicios de activos virtuales (VASP) que cumplan por completo con todas las limitaciones de divisas y controles de capital aplicables para mitigar el riesgo de sustitución de divisas (“Economics and the Libra

Reserve”, s.f., párrafo 6). Además, tienen la expectativa de que cada SCS estará respaldada por una red competitiva de revendedores e intercambiadores que garanticen seguridad sobre la convertibilidad de las monedas Libra en su moneda local.

Muy relacionado con lo explicado en apartados anteriores, sobre la posibilidad de que los bancos centrales emitan su propia moneda digital (CBDC), la red Libra tiene la intención de integrar las CBDC para así reducir los riesgos asociados a la gestión de la reserva (“Economics and the Libra Reserve”, s.f., párrafo 7).

Paralelamente a esto, Libra tenía como plan lanzar una billetera digital llamada Calibra que sería como muestra del funcionamiento de la tecnología Libra, permitiendo a cualquiera con un móvil inteligente adquirir y almacenar la moneda digital Libra y pagar varios bienes con ella. Esta billetera digital ahora soportará múltiples monedas, entre las que está Libra. Esta billetera en vez de estar disponible a nivel mundial estará restringida a aquellos países cuyos gobiernos usan monedas soportadas por el proyecto Libra. (Facebook is shifting its Libra cryptocurrency plans after intense regulatory pressure”, citando a “Facebook Revamps Libra Plans, Bowing to Regulators”, 3 de marzo de 2020, párrafo 3).

Finalmente, cabe comentar que para garantizar la protección del consumidor, Libra está desarrollando un regulador de capital para protegerse de potenciales pérdidas de crédito o del riesgo operacional del sistema de pagos de Libra, como por ejemplo fallos en el sistema o fraude interno o externo.

En conclusión, pese a que Libra ha sufrido muchos cambios respecto a su concepción inicial, que ha sido explicada al inicio de este trabajo para que el lector sea consciente de cuán difícil es conseguir algo como lo que pretendía Facebook por las implicaciones financieras que esto tiene, su nueva red, que va a permitir el uso de otras monedas digitales que representan monedas nacionales en su forma digital, puede hacer que su uso a nivel internacional se vea incrementado por dos razones: la primera es que aquellos que pudieran desconfiar de la moneda Libra tienen otras alternativas como son las SCS y el segundo es que, debido a que ahora “acoge” otras monedas nacionales que podrás ser usadas por sus usuarios, esto por un lado hace que los bancos centrales puedan tener más control sobre Libra y por otro lado hace que la oposición de gobernantes de distintos países disminuya.

Libra se ha dado de frente ante la realidad de que no se quiere una moneda digital que pueda poner en peligro la estabilidad financiera de los países, pero con su cambio de estructura puede que consiga superar todas las barreras que ha tenido hasta la fecha, esto es algo que falta por ver.

Hay otras compañías que trataron de solucionar el problema de los bajos niveles de bancarización en los países poco desarrollados y los altos costes asociados al envío de remesas a nivel local, y han sido exitosas en su cometido. Este es el caso de M-Pesa.

10 M-Pesa, un caso de éxito

M-Pesa es un producto de telefonía móvil lanzado en 2007 por Vodafone, con la ayuda de Safaricom y Vodacom. Safaricom es una compañía de telecomunicaciones de Kenia, compartida por Vodafone y el gobierno de Kenia (“Vodafone targets Africa’s unbanked with ambitious

plans for M-Pesa”, 18 de diciembre de 2019, párrafo 8). Por otro lado Vodacom es una empresa de telecomunicaciones Sud Africana que trabaja principalmente en Tanzania, República Democrática del Congo, Mozambique y Lesotho. Permite realizar pagos con el teléfono móvil, enviar y recibir dinero entre usuarios y no usuarios, retirar dinero efectivo en cajeros y otros muchos servicios. Es un servicio de banca sin sucursales, por lo que permite a los usuarios la realización de transacciones sin la necesidad de ir a una sucursal bancaria (“M-Pesa”, s.f.).

En lugar de que tengas que descargar una aplicación, M-Pesa utiliza tu teléfono móvil como cartera, convirtiendo tu línea telefónica y tu tarjeta SIM en una cuenta bancaria. Mandar dinero es el equivalente a mandar un mensaje. El teléfono crea algo llamado “código rápido” y envía una solicitud mediante los ordenadores del operador. Así es como se confirma la transacción (“M-PESA: Cómo Kenia ha revolucionado los pagos móviles”, 30 de abril de 2018, párrafos 3-4).

Aunque en un principio tenía como objetivo crear un servicio que permitiese a los beneficiarios de microcréditos recibir el dinero y pagar los préstamos utilizando una red de revendedores de recargas de móvil de Safaricom, reduciendo así los costes y permitiendo a las instituciones de microfinanzas ofrecer tasas más competitivas a los usuarios, este objetivo se redefinió y pasó a ser tener como objetivo el envío de remesas y el uso como medio de pagos (“M-Pesa”, s.f.), viendo su oportunidad de desarrollo en los países menos desarrollados donde Libra también pretende ayudar a mejorar el acceso a servicios financieros de la población. Es debido a este cambio de negocio que pasó a implantarse en Kenia en 2007, llegando a tener 14 millones de usuarios y casi 30000 distribuidores en este país solo cuatro años de iniciarse en el mismo, y consiguiendo que millones de personas sin cuenta bancaria pudieran realizar transferencias de dinero.

Una de las razones por las que ha tenido tanto éxito en Kenia es debido a que, tal y como se indica en (“M-PESA: Cómo Kenia ha revolucionado los pagos móviles”, 30 de abril de 2018, párrafos 8-9), muchos de los kenianos que trabajan en las ciudades tienen parientes en zonas rurales y resultado difícil acceder a ellas debido a la falta de infraestructuras para enviar dinero efectivo por correo postal.

Este país, al igual que otros muchos países de África, se caracteriza por una infraestructura bancaria muy limitada. Es por esto que M-Pesa supuso un cambio tan importante, ya que ahora, mediante el uso exclusivo del teléfono móvil, los keniatas podían enviar dinero a sus amigos, familiares o clientes apretando unos botones, sin necesidad de tener que depender de una entidad bancaria. Según estimaciones, una cuarta parte del producto interior bruto de Kenia es procesado en esta plataforma. Además, es usado para pagar salarios, liquidar facturas, para conceder pequeños préstamos, en seguros y para muchos otros tipos de servicios (“Vodafone targets Africa’s unbanked with ambitious plans for M-Pesa”, 18 de diciembre de 2019, párrafo 19-20). Por lo tanto, lejos de ser una plataforma cuya única utilidad está reducida a la realización de transferencias, su integración en la economía de Kenia es mucho más importante debido al reemplazamiento de los móviles básicos por *smartphones*.

Desde entonces se ha implementado en otros países como Kenia (2007), Tanzania(2008), Lesotho (junio de 2013), República Democrática del Congo (?), Ghana(2015), Mozambique (mayo de 2013) y Egipto (julio de 2013), convirtiéndose así en la mayor plataforma de pagos de África, con 40 millones de usuarios y más de mil millones de transacciones mensuales (“Vodacom and Safaricom joint venture to accelerate M-Pesa expansion”, 6 de abril de 2020, párrafo 6).1

No todo ha sido un éxito. En países como India, Afghanistan, Rumanía, Albania y Sudáfrica, donde han intentado dar su servicio han fracasado y han cerrado el negocio allí.

Con tal de estudiar que impacto ha tenido M-Pesa en estos países, cual es la situación actual de los países en desarrollo respecto a los desarrollados en cuanto a nivel de bancarización y con tal de estudiar los costes asociados al envío de remesas a nivel internacional, se hará un análisis estadístico propio a partir de diversas bases de datos. Cabe concluir cuan necesario es que haya otras alternativas como Libra para intentar mejorar la situación de estos países e intentar disminuir los costes de las transferencias internacionales.

Con este análisis se pretende comprender cuan necesarias son las alternativas a los métodos de pagos convencionales y como éstos últimos han evolucionado a lo largo de la última década. La evolución que los métodos convencionales tengan puede tener implicaciones importantes en las posibilidades de éxito de proyectos como Libra, ya que en caso de que los costes del envío de dinero a nivel internacional se reduzcan lo suficiente esto puede neutralizarla parcialmente.

11 Análisis estadístico de las remesas y su evolución

Para el análisis estadística que se hará a continuación se utilizarán tres bases de datos para cruzar los datos y comparar así los resultados:

- Costes asociados al envío de remesas a nivel mundial con datos de 2011 a 2019, con las salidas y entradas de remesas de casi todos los países del mundo. **Estos envíos de remesas registrados no son a nivel doméstico, sino a nivel internacional.** Estos datos se obtienen a partir de la base de datos del banco mundial (The World Bank, Remittance Prices Worldwide, disponible en <http://remittanceprices.worldbank.org>). Esta base de datos consta de 365 “corredores de países” de todo el mundo. Contiene flujos de remesas que se envían de 48 países y son recibidos en 105 países. Los datos han sido capturados en los lugares desde los que se envían/reciben más remesas en cada país. Los datos son recogidos por investigadores que actúan como si fueran clientes contactando a las firmas, los cuales, dentro de cada corredor, han recogido los datos en el mismo día para controlar las fluctuaciones en los tipos de cambio y otros cambios en las estructuras de tarifas. (“REMITTANCE PRICES WORLDWIDE Methodology”, s.f., párrafos 3-4). De esta base de datos se utilizarán las siguientes variables:
 - Origen: país desde el que se envía el dinero.
 - Destino: país al que se envía el dinero.
 - Tipo de firma: tipo de proveedor de servicios de remesas que ofrece el servicio. Las empresas encuestadas dentro de cada corredor se seleccionan para cubrir la máxima cuota de mercado de remesas posible, con el objetivo de una cuota de mercado agregada mínima del 80 por ciento. Los principales que hay son los siguientes:
 - * Operador de Transferencia de Dinero: Estos son facilitadores del pago de transacciones a nivel internacional, cuya definición en (IMF 2009) es la de “compañías financieras (pero generalmente no los bancos) que participan en la transferencia transfronteriza de fondos, utilizando su sistema interno o el acceso a otra red bancaria transfronteriza”.
 - * Banco
 - * oficina postal
 - * Cooperativa de ahorro y crédito: es un tipo de cooperativa financiera que brinda servicios bancarios tradicionales (“Credit Union”, 16 de mayo de 2019, párrafo 1).
 - cc(x) cantidad de denominación: cantidad de dinero enviada en dólares. Solo hay dos cantidades recogidas que son 200 y 500 dólares. La cantidad de remesas que será estudiada será la de 200 dólares. Este coste suele representar el cargo que paga el remitente en el punto de inicio, que generalmente varía con la cantidad enviada. En algunos casos, puede haber tarifas e impuestos cobrados en el destino que no se han detectado en esta base de datos (“REMITTANCE PRICES WORLDWIDE Methodology”, s.f., párrafo 7).

- cc(x) coste total %: coste total de la transacción respecto a la cantidad enviada.
- La base de datos Global Findex, que contiene el conjunto de datos más completo del mundo sobre cómo los adultos ahorran, piden prestado, realizan pagos y gestionan el riesgo. En esta base de datos hay 776 variables recogidas en tres años: en 2011, 2014 y 2017. Algunas de ellas solo están disponibles para 2014 y 2017. Aunque se podría estudiar con mucha profundidad se escogen las variables que se consideran más relevantes teniendo en cuenta el tema tratado en este trabajo. Se utilizará la clasificación que ésta hace de los países, que es la siguiente:
 - África subsahariana (excluyendo altos ingresos)
 - América Latina & el Caribe (excluyendo altos ingresos)
 - Asia del Sur
 - Asia Oriental & Pacífico (excluyendo altos ingresos)
 - Economías de altos ingresos
 - Europa & Asia Central (excluyendo altos ingresos)
 - Oriente Medio & Norte de África (excluyendo altos ingresos)
- Datos de las entradas y salidas de remesas a nivel de país entre los años 2000 y 2019 para así poder ver como ha estado evolucionando en términos nominales en las diferentes regiones.

12 Análisis de la entrada y salida de remesas

El Banco Mundial define las remesas como la suma de:

- las transferencias personales, que incluyen:

Todas las transferencias actuales en efectivo o en especie entre individuos residentes y no residentes, independientemente de la fuente de ingresos del remitente (e independientemente de si el remitente recibe ingresos de mano de obra, de emprendimiento o de propiedad, beneficios sociales y cualquier otro tipo de transferencias o disposición de activos) y la relación entre los hogares (independientemente de si son personas relacionadas o no).

- la compensación de los empleados, que se refiere a:

Los ingresos de los trabajadores fronterizos, estacionales y otros trabajadores a corto plazo que están empleados en una economía donde no son residentes y de los residentes empleados por entidades no residentes. La remuneración de los empleados representa una remuneración a cambio del aporte laboral al proceso de producción aportado por un individuo en una relación empleador-empleado con la empresa. La remuneración de los empleados tiene tres componentes principales: sueldos y salarios en efectivo, sueldos y salarios en especie y contribuciones sociales de los empleadores. La remuneración de los empleados se registra bruta e incluye los montos pagados por el empleado como impuestos o para otros fines en la economía donde se realiza el trabajo.

En este capítulo se analiza la evolución de las remesas según la región de la que proceden y la región a la que se envían. Además, se muestran cuales son los países que envían y reciben más remesas. Con esto se pretende ver cual es la situación actual y como ha sido la evolución de una de las variables que está detrás de la voluntad de crear Libra.

Cabe destacar que los países de economías de altos ingresos desde los cuales se envían remesas recogidos en la base de datos son Estonia, Hungría, República de Corea, Letonia, Lituania y Polonia. Los países de economías de altos ingresos son el resto de países con alto nivel de ingresos como por ejemplo Alemania, Francia, España, Estados Unidos, Reino Unido, etc. .

En la figura 9 hay cuatro gráficos. Los que están a la izquierda informan sobre la evolución de las remesas en valor nominal (millones de dólares estadounidenses) entre los años 2000 y 2019, según la región desde la que se envían (gráfico superior) o la región que los recibe (gráfico inferior).

Se observa como en el caso de entrada de remesas todas las regiones tienen una tendencia creciente, así como que se está produciendo una convergencia entre las economías de altos ingresos, Asia Oriental & Pacífico y Asia del Sur. Respecto a las salidas de remesas la imagen es bastante diferente, ya que las economías de altos ingresos están muy lejos del resto de regiones. Esto es debido a dos factores: el primero es que este conjunto de países de las economías de altos ingresos es superior al del resto de regiones, el segundo es que la población de países pobres, cuando emigra, acostumbra a ir a países con alto nivel de renta en los que puedan tener la oportunidad de ganar un buen sueldo y así poder enviar dinero a los familiares de su país de origen.

Solo con este análisis se hubiera cometido el error de no analizar cual ha sido la evolución respecto

al año 2000, es por esto que usan números índice con base 100 en el año 2000 tanto para la entrada de remesas como salida. En cuanto a la entrada de remesas, los países de África subsahariana y Asia Oriental & Pacífico son los que han tenido mayor crecimiento, teniendo en el año 2019 una entrada de remesas un 1100% superior a la del año 2000. Por otro lado, los países de las economías de altos ingresos son los que tienen menos crecimiento de entrada de remesas. Aunque el causante de esto es multifactorial y está fuera de lo que se pretende estudiar, lo que es seguro es que uno de los causantes que hay detrás de esto es la bancarización de los países menos desarrollados, que como se verá más adelante, ha sido bastante importante y ha permitido a millones de personas que no tenían acceso a este servicio tan esencial el poder recibir dinero de familiares que están en el extranjero.

En cuanto a la salida de remesas, en África subsahariana, en las economías de altos ingresos y en América Latina & el Caribe, el crecimiento en números índice de la salida de remesas no ha sido muy importante en comparación con el resto de agrupaciones. En estas otras agrupaciones hay un crecimiento hasta el año 2013, año en el que todas, a excepción de Asia Oriental y Pacífico tienen una bajada importante.

La bajada de Europa & Asia Central está debida principalmente a la Federación Rusa, que pasó de tener 32640 millones de US\$ en salidas de remesas en 2014 a 19688 millones en 2015. Esto fue causado principalmente por que, tal y como se comenta en ("How to pay for purchases in Russia without losing money in ruble exchange?", s.f., párrafos 7- 8), el rublo ha sufrido una fuerte devaluación en los últimos años. En noviembre de 2014, el Banco Central de Rusia (BCR) dejó el rublo en flotación libre, eliminó la banda del tipo de cambio y terminó sus intervenciones regulares en el mercado de divisas, para así evitar devaluaciones. Mientras que en octubre de 2014 estaba situado en unos 50 rublos por dólar, el rublo se ha devaluado hasta llegar a 90 rublos por dólar a principios de 2016. De 2016 a 2018 se ha apreciado, estando alrededor de los 70 rublos por dólar. Esto explica por qué tiene esta bajada tan importante y a partir de 2016 empieza a subir.

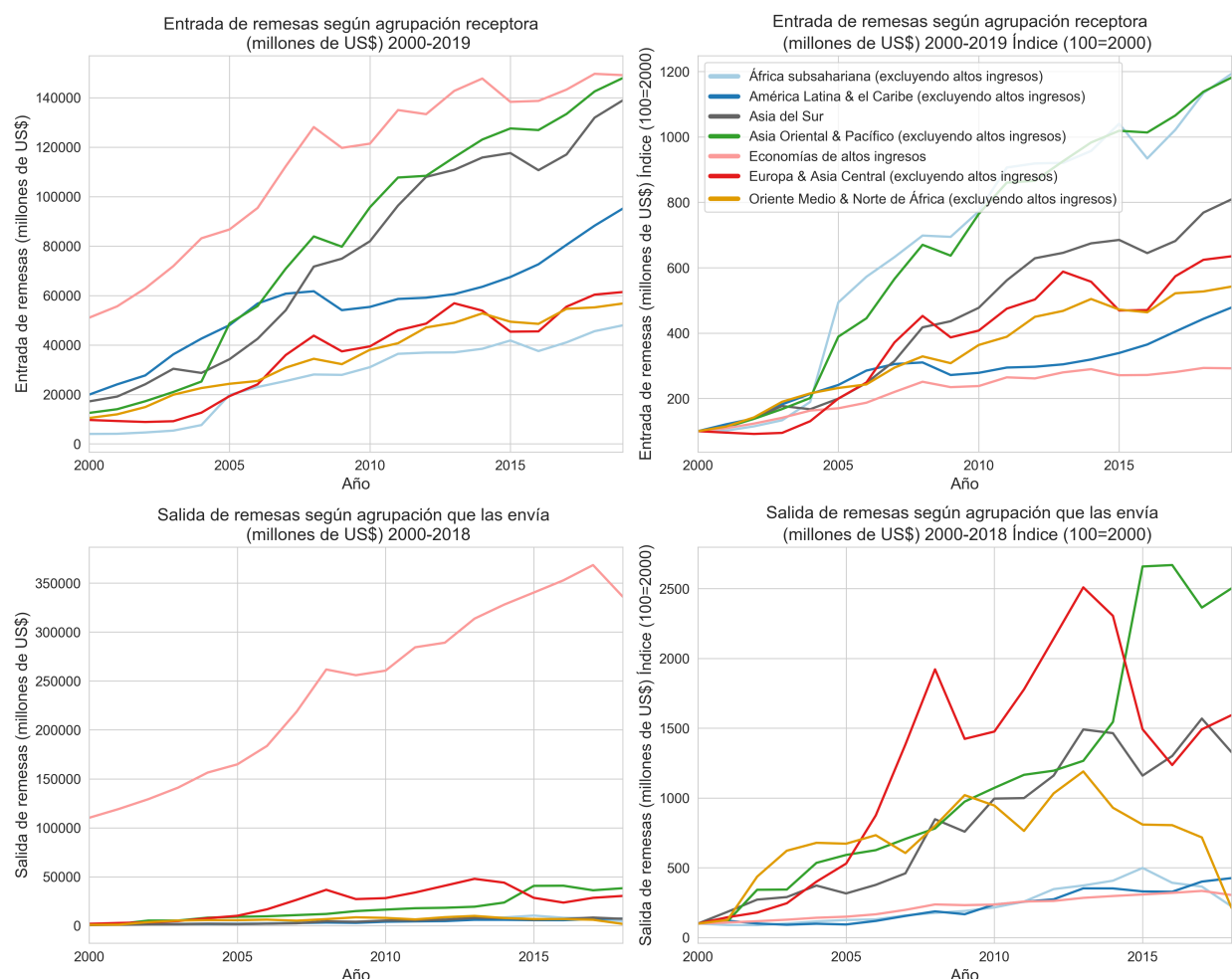


Fig. 9: Total de remesas (millones de dólares estadounidenses) recibidas (gráficos de arriba) y enviadas (gráficos de abajo) por 7 agrupaciones de países del año 2000 al año 2019. Elaborado a partir de: The World Bank, Remittance Prices Worldwide, disponible en <http://remittanceprices.worldbank.org>

Una vez analizado la evolución atendiendo a agrupaciones de países, sería interesante analizarlo por país para así tener una idea de cuales son los países con mayor relevancia a nivel mundial en cuanto a entrada y salida de remesas.

En la figura 10 se observan los 10 países con mayor entrada de remesas en millones de US\$. En los tres primeros puestos están India, China y México que, para el año 2019, tienen unas salidas de 82202 millones, 70265 millones y 38655 millones respectivamente. India recibe las remesas principalmente de los Emiratos Árabes, Estados Unidos y Arabia Saudita mientras que China de Estados Unidos, Hong Kong y Japón (“China - Remesas de migrantes” e “India - Remesas de migrantes”, s.f, párrafo 4).

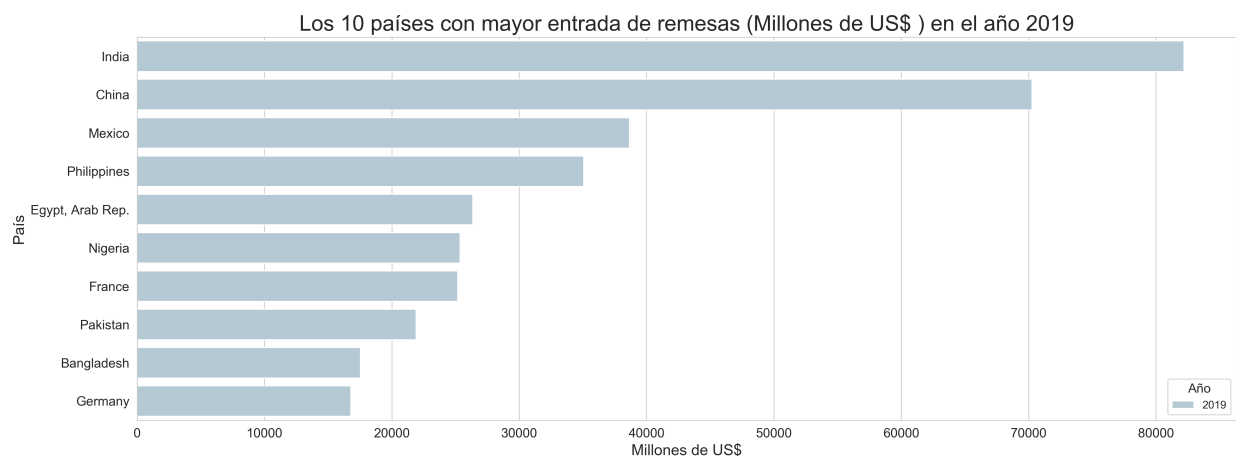


Fig. 10: Los 10 países con mayor entrada de remesas (millones de dólares estadounidenses) en el año 2019. Elaborado a partir de: The World Bank, Migration and Remittances Data

En la figura 11 se observan los países con mayor envío de remesas para el año 2018, último disponible. Los dos primeros son Estados Unidos y Arabia Saudí.

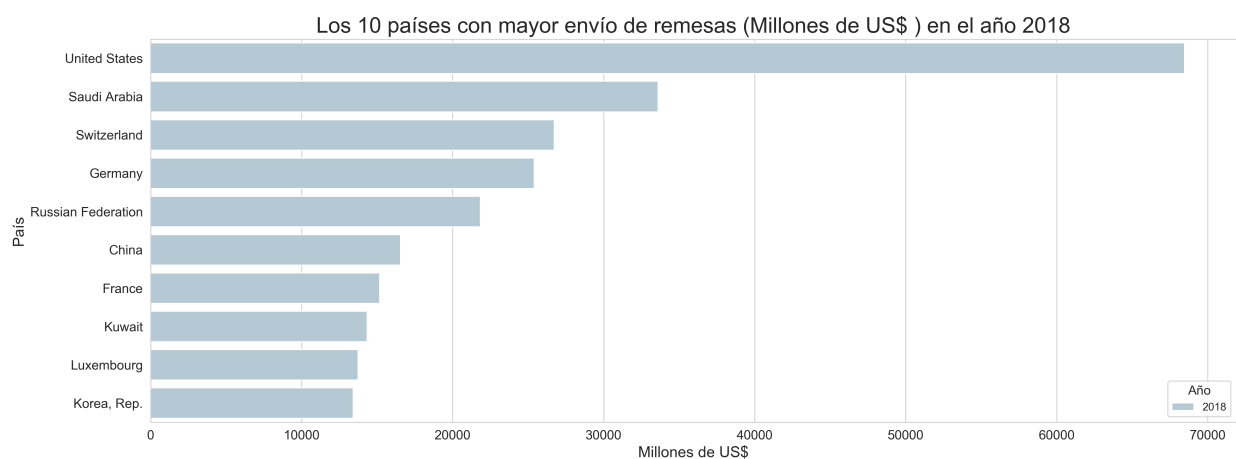


Fig. 11: Los 10 países con mayor entrada de remesas (millones de dólares estadounidenses) en el año 2018. Elaborado a partir de: The World Bank, Migration and Remittances Data

13 Análisis coste (%) de enviar 200 dólares

Una vez analizada cual ha sido la evolución de las remesas a nivel mundial hace falta poner el foco en el coste asociado a enviarlas, ya que éste es uno de los motivos por los que Libra quiere implementarse a nivel mundial. Países sobre todo de África sufren de altos costes y es por ello que Libra quería ser una alternativa menos costosa para aquellos que quieren enviar remesas a sus familiares. Todas las figuras tienen unos colores consistentes con los del apartado anterior para así facilitar su análisis al lector.

La figura 12 tiene la misma estructura que la figura 9, solo que en vez de representar la evolución de las remesas, representa el coste promedio de las transacciones, que suele estar a cargo del remitente, y que está representado en términos porcentuales respecto a la cantidad enviada. Por lo tanto, en esta figura está representado el coste promedio asociado a enviar remesas según la región que las recibe (gráficos superiores) y el coste promedio asociado a enviar desde cada una de las agrupaciones (gráficos inferiores).

Se puede ver como recibir remesas en África subsahariana tiene un coste muy superior al del resto de regiones. Lo mismo sucede con el envío de remesas desde África. Cabe destacar como el coste asociado a recibir remesas (gráfico izquierda-superior) ha tenido un decrecimiento muy importante en todas las agrupaciones a lo largo del periodo 2011-2019, destacando África subsahariana, que ha pasado de tener un coste cercano al 13% en el año 2000 a tener uno próximo al 9% en el año 2019.

Si se analiza mediante los números índice, se observa como todas las regiones, a excepción de Asia Oriental & Pacífico y Europa y Asia Central han tenido un decremento del coste asociado a enviar 200 dólares estadounidenses de remesas desde las mismas. Por otro lado, si se pone el foco en la región a la que van destinadas las remesas, el coste ha disminuido de una manera muy importante en todos los destinos, teniendo la mayor reducción cuando el destino de las remesas es África subsahariana o las economías con altos ingresos.

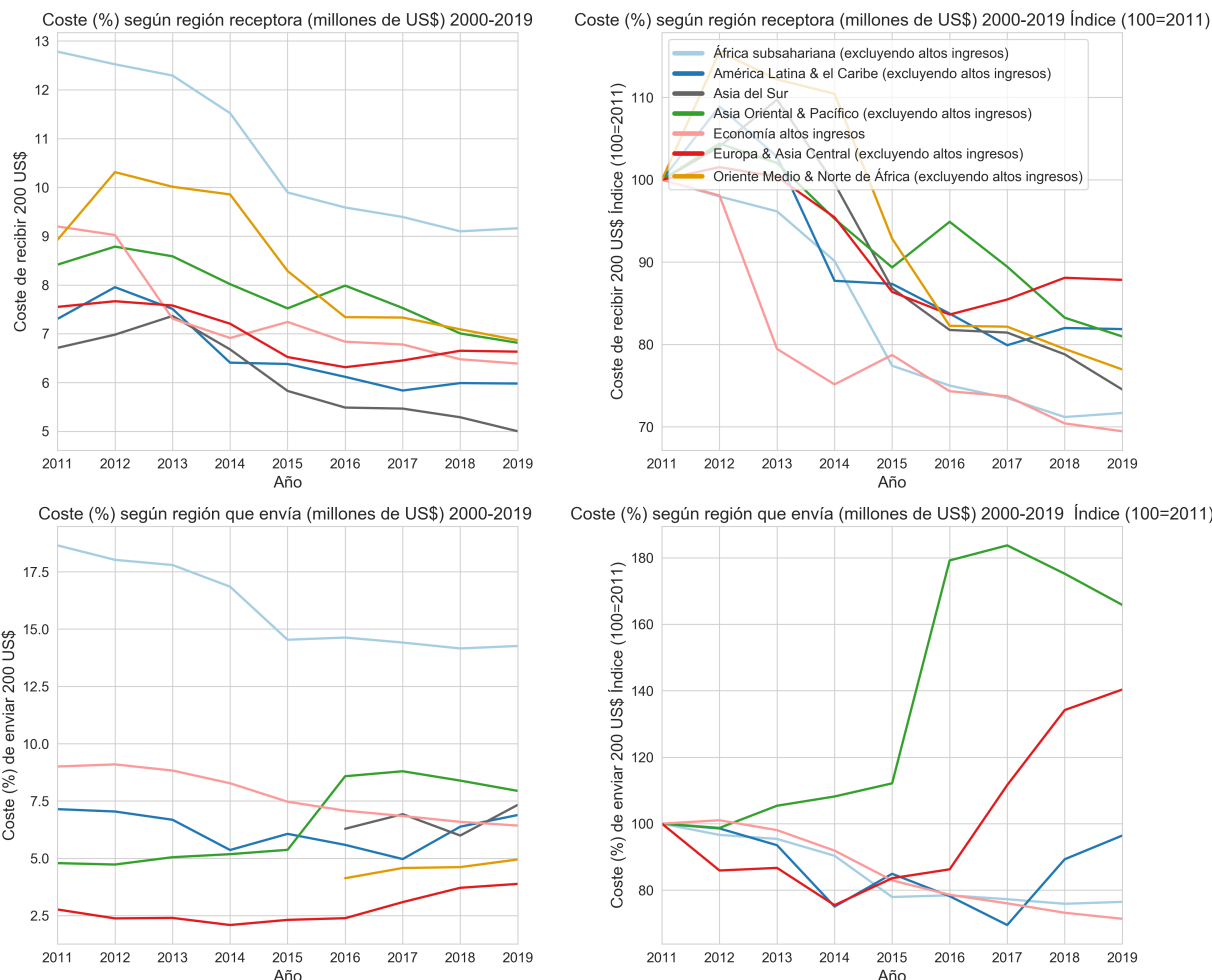


Fig. 12: Coste promedio de enviar remesas del año 2000 al año 2019 según la agrupación de países que las recibe (gráficos de arriba) o las agrupaciones que las envían (gráficos de abajo). Elaborado a partir de: The World Bank, Remittance Prices Worldwide, disponible en <http://remittanceprices.worldbank.org>

Como se ha comentado anteriormente, los datos se han recogido en diferentes firmas como son los bancos, operadores de transferencia de dinero u oficinas de correo, por lo que cabe estudiar como han evolucionado los costes de enviar remesas según la firma que ha sido utilizada.

En la figura 13 están los costes promedios asociados al envío de remesas según la institución que se ha utilizado. Respecto al año 2011, todas las instituciones menos las instituciones financieras no bancarias han reducido su coste. Las que tienen una mayor reducción son las que son Bancos / Operadores de transferencia de dinero que han pasado de un coste cercano a 18 a un coste de 12.

Los bancos y los operadores de transferencia, que son los que representan más de un 90% de las instituciones utilizadas en la base de datos del banco mundial, siendo el segundo el más común, representando un 78%, y por lo tanto los que deberían analizarse con más detalle. Ambos han

tenido una reducción parecida en términos relativos, pasando de un precio medio en 2011 de 12.82 y 7.22 a uno de 10.36 y 6.02 en 2019 respectivamente.

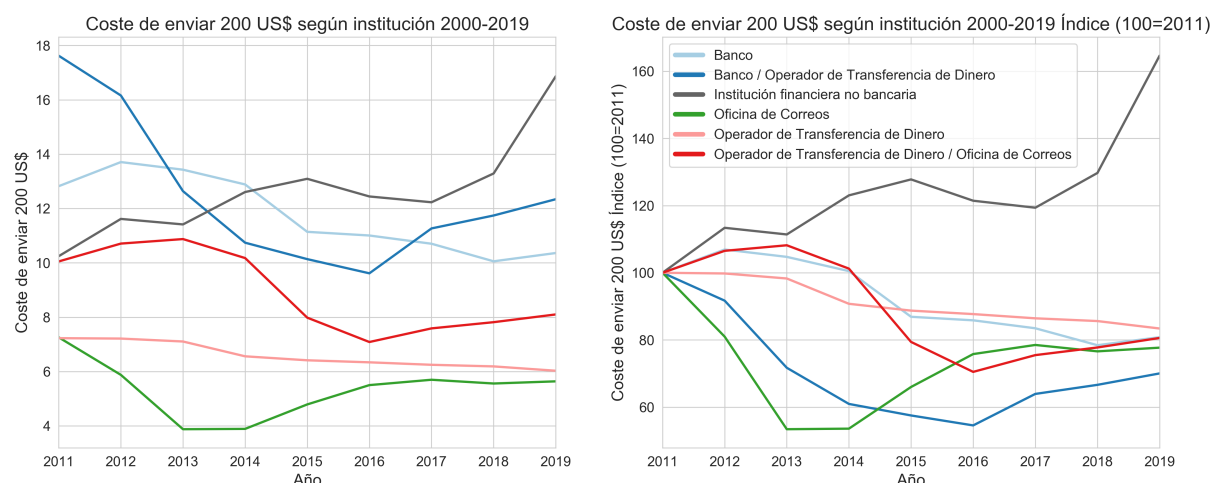


Fig. 13: Coste promedio de enviar remesas según la institución encargada de hacer la transferencia del año 2011 al 2019. Elaborado a partir de: The World Bank, Remittance Prices Worldwide, disponible en <http://remittanceprices.worldbank.org>

Tal y como se menciona en (“How Is the International Money Transfer Market Evolving?”, s.f., párrafo 7), los costes de enviar remesas han disminuido debido a una competencia más saludable ya la presión del G20 y la ONU, que tienen como objetivo un coste promedio para el consumidor del 5% y 3% respectivamente. Estos organismos prestan mucha atención al mercado de servicios de remesas debido a su popularidad entre los trabajadores de bajos salarios y al papel que juegan las remesas en el desarrollo económico de los mercados emergentes. Los bancos, tal y como se menciona en este artículo, enfrentan la mayor presión de este impulso para reducir las tarifas, ya que son los proveedores de remesas más caros tal y como se ha visto confirmado en la figura 13, en la que aunque se ve una bajada de su coste, sigue siendo bastante más elevado que el de los Operadores de Transferencia de Dinero.

Una vez se ha analizado la evolución del coste según la institución, se pone el foco en analizar como ha variado el coste según la región receptora de las remesas y analizando solo los bancos y las operadoras de transferencia de dinero, que son las principales utilizadas. Además, para el resto de instituciones hay muchos datos faltantes debido a la baja frecuencia de uso que tienen en comparación con las dos mencionadas.

En la tabla 1 se muestra la evolución del año 2011 al 2013 del coste promedio (%) de enviar remesas según si se utiliza un banco o un operador de transferencia de dinero y según la región receptora de las remesas. En el caso de África subsahariana, si se pone el foco en la banca, es la región que más sufre de altos costes, aunque como se puede ver éste ha tenido una reducción entre los años 2011 y 2015 pero en el año 2019 ha vuelto a incrementarse, estando en 16.92%, muy por encima del resto de agrupaciones de países. Contrariamente a esto, con el Operador de Transferencia de Dinero, en el caso de África Subsahariana se observan tres cosas:

- En primer lugar el coste es muy inferior al de la banca, es por esto que tiene un mayor uso, tal y como se había comentado anteriormente.
- En segundo lugar del 2011 al 2019 ha habido una reducción de más de casi dos puntos porcentuales, pasando de 9.67% a 7.86%. Esta tendencia bajista también se observa en el resto de agrupaciones a excepción de Europa & Asia Central.
- En último lugar cabe mencionar es que, pese a que estos costes son inferiores a los de la banca, siguen siendo más elevados que los del resto de agrupaciones.

Coste promedio (%) de enviar remesas según institución y región receptora			
	2011	2015	2019
Banco	12.84	11.16	10.38
África subsahariana (excluyendo altos ingresos)	19.12	15.34	16.92
América Latina & el Caribe (excluyendo altos ingresos)	8.47	7.82	7.39
Asia del Sur	9.16	9.74	8
Asia Oriental & Pacífico (excluyendo altos ingresos)	13.15	11.58	11.15
Economía altos ingresos	21.3	2.12	2.9
Europa & Asia Central (excluyendo altos ingresos)	12.08	8.66	8.28
Oriente Medio & Norte de África (excluyendo altos ingresos)	11.2	11.72	8.9
Operador de Transferencia de Dinero	7.22	6.41	6.03
África subsahariana (excluyendo altos ingresos)	9.67	8.27	7.86
América Latina & el Caribe (excluyendo altos ingresos)	6.97	6.15	5.9
Asia del Sur	5.34	4.78	4.3
Asia Oriental & Pacífico (excluyendo altos ingresos)	7.62	6.58	5.92
Economía altos ingresos	7.11	7.59	6.54
Europa & Asia Central (excluyendo altos ingresos)	6.03	5.91	6.34
Oriente Medio & Norte de África (excluyendo altos ingresos)	6.9	6.45	6.24
Total general	9.1	7.55	6.8

Tabla 1: Coste promedio de enviar remesas mediante el uso de un banco o un Operador de Transferencia de Dinero para los años 2011, 2015 y 2019 según la agrupación de países que las recibe. Elaborado a partir de: The World Bank, Remittance Prices Worldwide, disponible en <http://remittanceprices.worldbank.org>

14 Análisis base de datos Findex

Hasta ahora el análisis que se ha hecho ha sido a nivel global. Se ha visto como, pese a que los costes asociados al envío de remesas tienen una tendencia descendiente, estos siguen siendo muy elevados en regiones como África subsahariana, por lo que empresas como Facebook podrían ser una alternativa más barata que podría ayudar a millones de personas a tener la posibilidad de reenviar o recibir remesas a un bajo coste.

Un segundo aspecto a estudiar es el nivel de acceso a servicios financieros, el uso del teléfono móvil y el porcentaje de la población que posee uno y cual ha sido su evolución en los últimos años, poniendo sobre todo el foco en África subsahariana, región que está en la peor situación. Es importante saber cual ha sido la evolución de estas variables, ya que es el bajo nivel de bancarización lo que justifica la aparición de otras alternativas para transferir dinero como lo es Libra y además es condición absolutamente necesaria que la población tenga teléfono móvil o acceso a internet porque sin esto no podrían acceder a ninguna billetera digital como Calibra.

En la figura 14 se muestra la evolución en los años 2011, 2014 y 2017 (los disponibles actualmente) de los adultos (%) con una cuenta en una institución financiera o que han utilizado un servicio monetario mediante un móvil en los últimos 12 meses. En ésta, además de estar las 7 agrupaciones de países que se han utilizado hasta ahora hay dos nuevas que son M-Pesa, países de África subsahariana en las que M-Pesa lleva años trabajando y No M-Pesa, países de África subsahariana donde M-Pesa no está presente. Esto se ha hecho con el objetivo de ver cual ha sido el efecto que ha tenido en estos países.

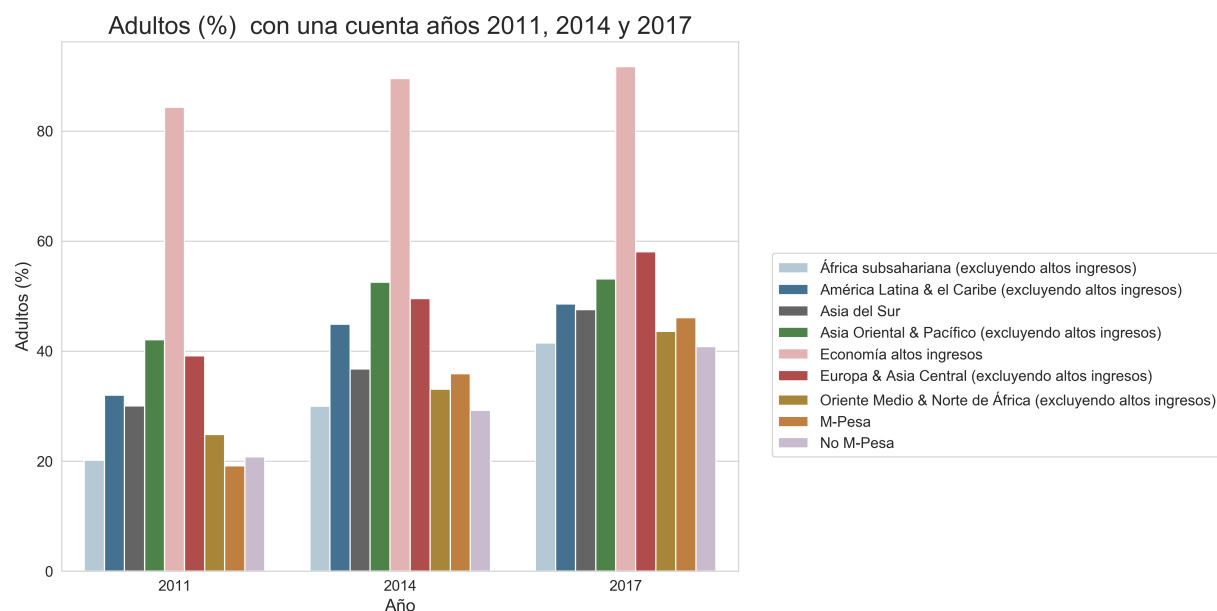


Fig. 14: Adultos que informan tener una cuenta (por sí mismos o junto con otra persona) en un banco u otro tipo de institución financiera o informan personalmente utilizando un servicio monetario mediante un móvil en los últimos 12 meses para los años 2011, 2014 y 2017. Elaborado a partir de: The Global Findex Database 2017

Las economías de altos ingresos destacan por tener un nivel de bancarización muy por encima del resto de regiones, del 91.67% en el año 2017, y con una tendencia ascendente. El resto de agrupaciones, pese a que han tenido una tendencia ascendente, siguen teniendo un porcentaje muy elevado de la población sin acceso a los servicios financieros. África subsahariana es la región que más destaca por el enorme crecimiento que ha tenido, pasando del 20% de adultos con acceso a una cuenta en 2011 al 40% en 2017. M-Pesa parece que ha tenido un efecto positivo pese a que en muchos de los países lleva pocos años establecida, ya que en el año 2011 la agrupación M-Pesa estaba por debajo de No M-Pesa y en los dos siguientes años registrados está por encima.

La figura 15 es más esclarecedora sobre cual ha sido el efecto de M-Pesa en África Subsahariana, donde todos los países han mejorado su situación (hay algunos para los que no se disponen datos). Kenia es el país en el que más tiempo lleva establecida y de donde, por lo tanto, se puede inferir con precaución debido a que cada país tiene sus circunstancias, cuales son los efectos a largo plazo de la implantación de este servicio. Como se puede ver, Kenia ha pasado de tener un 40% de adultos con acceso a una cuenta en 2011 a tener un 80% en 2017, destacando por encima del resto de regiones, las cuales están al rededor del 40% en 2017. Aunque es una realidad que cada país tiene su contexto social, político y económico y que por lo tanto puede haber otras razones detrás de cada porcentaje, está claro que M-Pesa, con sus más de 40 millones de usuarios en África subshariana, ha tenido un efecto positivo.

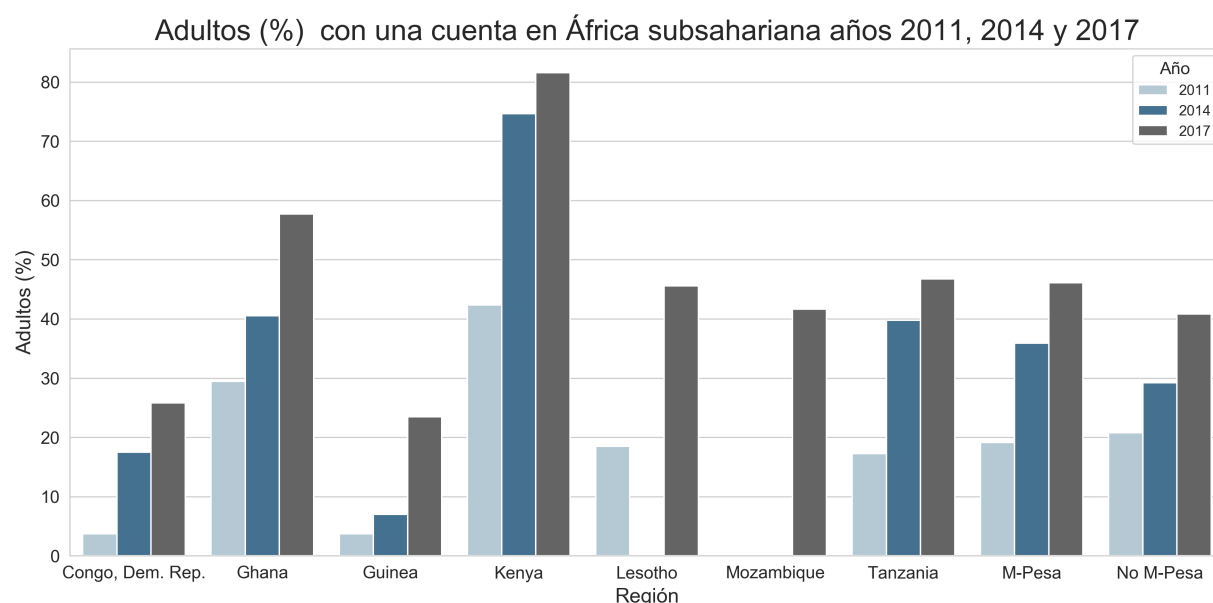


Fig. 15: Adultos de los países de África subsahariana que informan tener una cuenta (por sí mismos o junto con otra persona) en un banco u otro tipo de institución financiera o informan personalmente utilizando un servicio de dinero móvil en los últimos 12 meses para los años 2011, 2014 y 2017. Se han añadido dos agrupaciones nuevas que son M-Pesa, países que tienen en el servicio de M-Pesa y No M-Pesa, países que no disponen de este servicio. Elaborado a partir de: The Global Findex Database 2017

El porcentaje de adultos con teléfono móvil es otra variable que es clave a la hora de saber como

de viable es que Libra funcione como una alternativa para aquellos que no tienen acceso a los servicios financieros, ya que, tal y como hacen con M-Pesa, para enviar y recibir remesas a nivel regional, podrían hacer lo mismo a nivel internacional con Calibra. En la figura 16 se observa como, para el año 2017, único año disponible, todas las agrupaciones, incluido M-Pesa y No M-Pesa están por encima del 60%, estando las economías de altos ingresos en lo más alto con un 93.53%. Pese a que estos resultados son bastante elevados en comparación con el nivel de bancarización, cabría estudiar, en regiones como África subsahariana, cuantos de esos móviles tienen internet o la posibilidad de acceder a futuros servicios como Calibra debido al hecho de que podrían tratarse de móviles que solo permiten hacer llamadas.

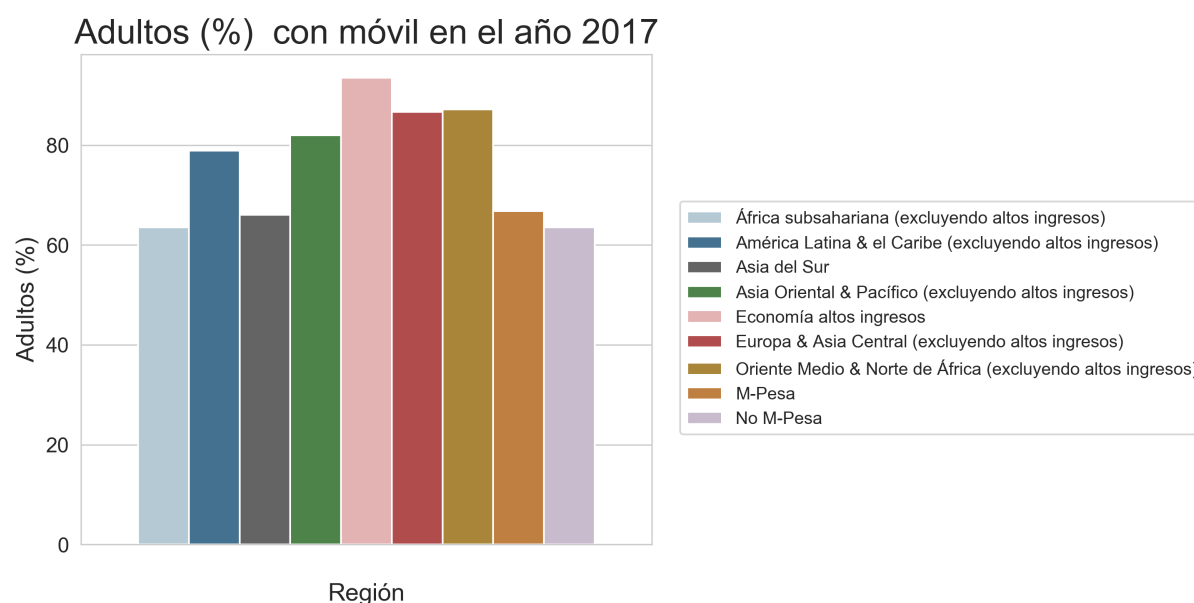


Fig. 16: Adultos que disponen de teléfono móvil en el año 2017 según agrupación de países a la que pertenecen. Se han añadido las dos agrupaciones extra de M-pesa y No M-pesa. Elaborado a partir de: The Global Findex Database 2017

Relacionado con esta última variable está el porcentaje de los adultos que usan un servicio monetario mediante el teléfono móvil (en las figuras 14 y 165 se veía reflejado esto junto a los adultos con una cuenta en un banco). En la figura 17 se puede ver como los países de África subsahariana destacan extremadamente por esto, sobre todo si se pone el foco en aquellos que tienen establecido el servicio M-Pesa. Seguido de esto están las economías de altos ingresos que del año 2014 al 2017 han tenido un incremento en término promedio muy importante, pasando del 5.6% al 16.51%. El resto de agrupaciones de países también han tenido un crecimiento de esta variable, por lo que se podría afirmar que esta tendencia ascendente del uso de servicios monetarios a través del teléfono móvil es algo generalizado a nivel mundial, pero en lo que destaca África subsahariana.

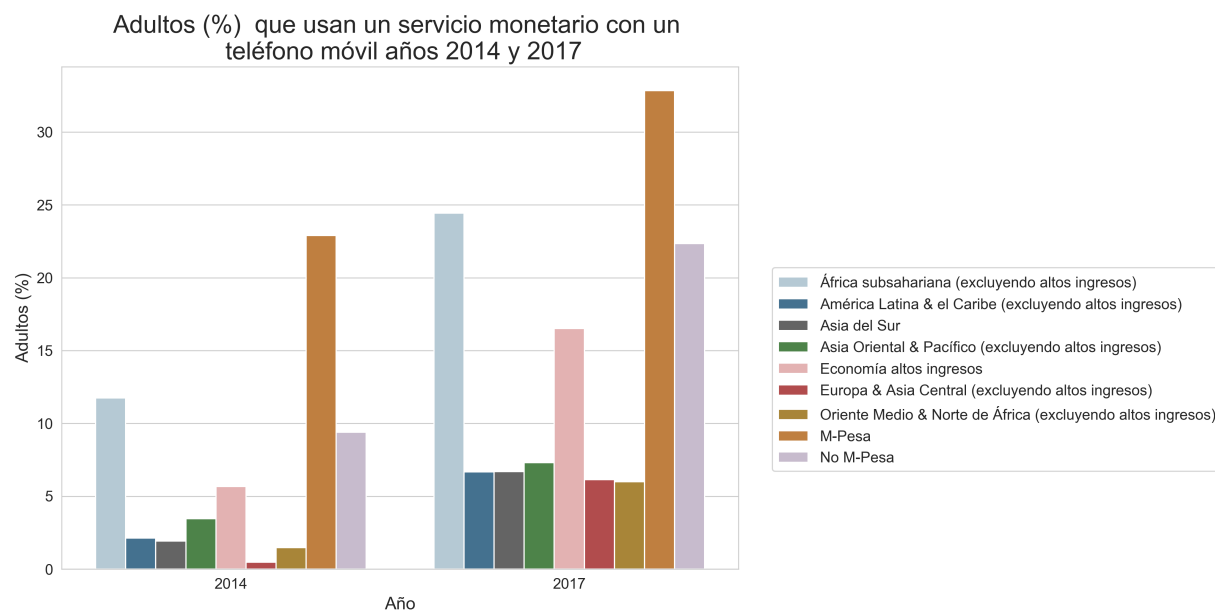


Fig. 17: Adultos que reportan haber utilizado un servicio monetario mediante un teléfono móvil para los años 2014 y 2017 según agrupación de países a la que pertenecen. Se han añadido las dos agrupaciones extra de M-pesa y No M-Pesa. Elaborado a partir de: The Global Findex Database 2017

Esta misma variable pero en África subsahariana se puede ver en la figura 18, siendo los resultados los mismos a los de la figura 15. Kenia se sitúa en niveles superiores al 70%, mientras que el resto países están por debajo del 40%. Como se puede ver, en los países para los que se disponen datos del 2014, todos han tenido un crecimiento de esta variable.

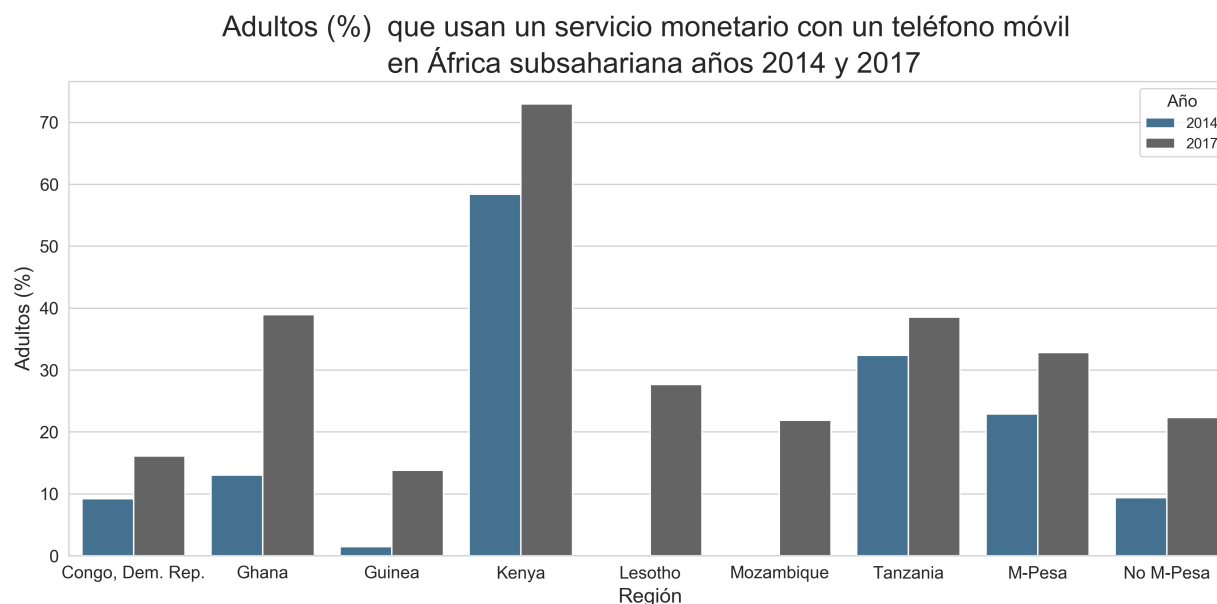


Fig. 18: Adultos que reportan haber utilizado un servicio monetario mediante un teléfono móvil para los años 2014 y 2017 según país de África al que pertenecen. Elaborado a partir de: The Global Findex Database 2017

La última variable que se va a proceder a analizar es los adultos (%) que han recibido o enviado remesas a nivel doméstico (dentro del país se encuentran la persona que envía y la persona que recibe las remesas). Como se ve en la figura 19, esto es un claro reflejo de por qué iniciativas como M-Pesa o Libra son tan necesarias. En África subsahariana un 24.43% de la población utilizó este servicio en el año 2017, mientras que el resto de agrupaciones de países están por debajo del 7%. Pese a ello, todas las agrupaciones de países han tenido un crecimiento de esta variable, en concordancia con lo que se ha visto en la figura 17. Esto seguramente es debido principalmente a tres factores:

- el primero es que debido a la poca infraestructura de transporte que hay en los países de África subsahariana. Si en una familia hay personas que trabajan a mucha distancia, por ejemplo en el campo, como pasa mucho en Kenia, este servicio es esencial, ya que no disponen de otra forma para enviar el dinero a sus familiares.
- el segundo motivo es la baja bancarización que hay, la cual les dificulta el poder ir a un banco a enviar dinero a familiares que viven en su mismo país pero a muchos kilómetros de distancia.
- el último factor es que a diferencia de lo que pasa en los países desarrollados, no hay grandes ciudades con millones de personas trabajando y en las que, por lo tanto, se puede acceder a cualquier servicio financiero.

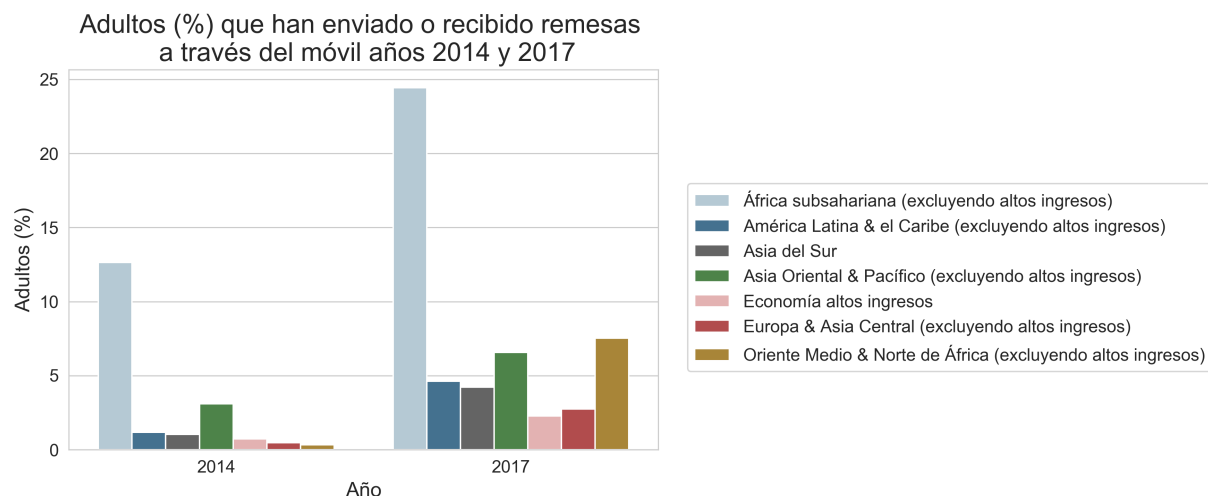


Fig. 19: Adultos que han enviado o recibido remesas a través del móvil en los años 2014 y 2017 según la agrupación de países a la que pertenecen. Elaborado a partir de: The Global Findex Database 2017

En la figura 20 se pone el foco en África subsahariana, donde los resultados siguen siendo consistentes con lo que se ha visto hasta ahora. En Kenia, donde hay 14 millones de usuarios de M-Pesa, el porcentaje de personas que han enviado o recibido remesas domésticas a través del móvil es de más del 60%, seguido de Ghana con el 38.94%. Por último cabe comentar que las regiones en las que está el servicio M-Pesa tienen un 10% en promedio de personas que realicen este servicio.

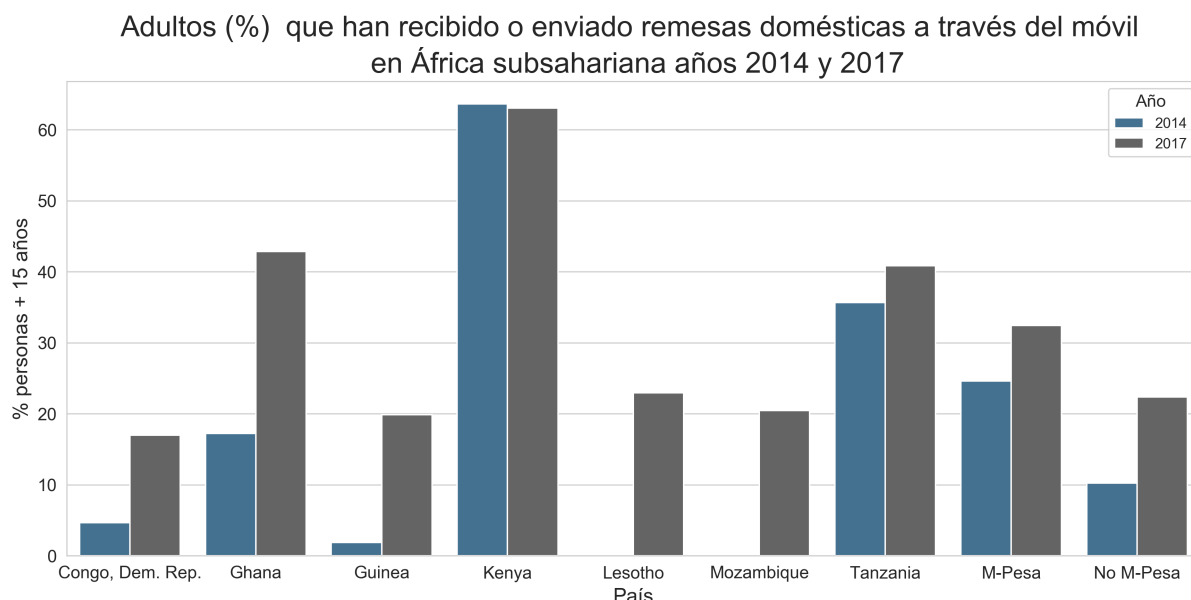


Fig. 20: Adultos que han enviado o recibido remesas a través del móvil en los años 2014 y 2017 según país de África subsahariana al que pertenecen. Elaborado a partir de: The Global Findex Database 2017

Cabe destacar que muy recientemente, en abril de 2019, ha salido una noticia de FOX News en la que se menciona un plan de Facebook muy ambicioso para tratar de solventar el problema de conectividad en África (“Facebook quiere resucitar a Simba en África, solo que el león se ha convertido en cable submarino”, 13 de abril de 2019, párrafos 1-5). Facebook planea construir un cable de fibra óptica alrededor de toda la costa africana, el cual recibirá el nombre de “Simba”. Aunque está en una fase muy temprana de su desarrollo, este proyecto ayudará a reducir los precios de ancho de banda y a la vez aumentará la accesibilidad a lo largo de todo el continente.

Paralelamente a esto, hay otros proyectos como el de Amazon que tienen la misma intención, pero en este caso planean hacerlo mediante satélites que proporcionen conexión a Internet de manera económica en todo el planeta.

Aunque estos proyectos pueden tardar muchos años en llevarse a cabo, pueden marcar una gran diferencia, ya que si se consigue que prácticamente la totalidad de la población africana tenga conexión a internet, esto podría hacer que el impacto de Libra sobre esta región fuera mucho mayor que si se implanta con la situación actual.

15 Análisis estadístico de las criptomonedas

A continuación se lleva a cabo un análisis estadístico de las criptomonedas situadas entre las 15 primeras en cuanto a capitalización (CoinMarketCap 2020) para las que se dispone de datos del 1 de febrero de 2017 al de diciembre de 2019. Primero se realizará un análisis exploratorio de las mismas en el que se evaluarán los principales estadísticos tanto para el precio de cierre que han tenido los últimos años como para la rentabilidad continua, definida de la siguiente manera:

$$r_i = \ln \left(\frac{P_i}{P_{i-1}} \right)$$

Siendo P_{i-1} el precio de cierre del día anterior al día i y r_i la rentabilidad continua del día i .

Una vez hecho el análisis exploratorio, se procederá a un análisis más exhaustivo de la rentabilidad continua. El precio es una variable que posee tendencia, y por lo tanto, al aplicar la transformación, lo que se pretende es que la variable transformada sea estacionaria en media y que tenga una varianza constante. Es debido a esto que en estudios parecidos a éste, el análisis del comportamiento de las criptomonedas se hace sobre la variable transformada.

El interés se pone en analizar el comportamiento distribucional de dicha variable para así poder inferir el riesgo asociado a cada criptomoneda a través de medidas como el CVaR o el VaR. **El cálculo de ambas medidas se hará a partir de la distribución que mejor ajuste proporcione para cada criptomoneda.**

Como se ha comentado en la metodología, a partir del análisis de que modelos de probabilidad son más idóneos para explicar el comportamiento de esta variable, se podrá concluir que distantes son las criptomonedas de la normalidad además. En vez de partir de la hipótesis de normalidad que nunca suele darse con datos financieros de alta frecuencia, se opta por estudiar varias de las distribuciones más comunes en finanzas. Con esto se conseguirá que el CVaR y el VaR calculados sean fiables, ya que, en caso de ser calculados a partir de una distribución paramétrica que no proporcione un buen ajuste, las conclusiones que se pudieran extraer sobre el riesgo con las medidas mencionadas no serían correctas. Con tal de asegurar que el ajuste es correcto, se utilizará el test de Kolmogorov Smirnov.

15.1 Introducción a las criptomonedas usadas

En total se han analizado seis criptomonedas, estas son: Bitcoin, Ethereum, Ethereum Classic, Ripple, Litecoin y Monero. Tether, como ya se ha comentado en la introducción, se trata de una criptomoneda estable, por lo tanto, pese a que cumple con la condición de estar entre las 15 primeras en capitalización (CoinMarketCap 2020) se descarta su análisis. Se procede a hacer una descripción con las características básicas de cada una de ellas:

Bitcoin: Se trata de la criptomoneda más popular hasta la fecha, fue creada en 2009, y tal y como se señala en (Bitcoin Project 2017), es la primera red descentralizada que usa la tecnología peer-to-peer o entre pares, la cual permite la gestión de transacciones y la creación de bitcoins de forma colectiva en una red sin el control de una autoridad central. Esta criptomoneda tiene una contabilidad pública llamada “blockchain”, la cual contiene todas las transacciones que

se producen y permite su validación. Son los usuarios, los cuales, mediante la resolución de un problema matemático, también llamado *proof of work*, añaden transacciones y a la vez son recompensados por ello.

Ethereum: Es la moneda alternativa al bitcoin más famosa y es más que una herramienta de pago o inversión. Ethereum se creó en 2015 y se basa en el uso de contratos inteligentes que permiten autoejecutar acciones de acuerdo a una serie de parámetros previamente programados. Vitalik, su creador, tenía la idea de hacer desaparecer al intermediario que está presente en muchas de las operaciones relacionadas con el mundo de la tecnología. Un ejemplo de contrato sería: “si la temperatura baja por debajo de 0 grados paga al granjero 3000 dólares” (Neel Mehta 2019). Estos contratos quedan registrados en una blockchain, donde en vez de depender de la interacción con un método centralizado de pago como las tarjetas de crédito, permite el pago mediante unas monedas llamadas Ether.

Ethereum Classic: Nace de un hackeo que se produjo a finales de julio de 2016 a la plataforma Ethereum donde se robaron 50 millones de Ethers. A partir de allí se crearon dos corrientes: los que querían modificar la Blockchain para devolver los Ethers robados y los que querían mantenerla. Los que decidieron manipularla crearon una nueva cadena, la hoy conocida como Ethereum y los que decidieron mantenerla igual crearon la plataforma Ethereum Classic (“Ethereum Classic (ETC): tutorial completo para dummies”, s.f., párrafos 3-5). Cabe destacar que Ethereum Classic desempeña las mismas funciones que Ethereum.

Ripple: Criptomoneda de la compañía estadounidense fundada en el 2012 bajo el nombre de Opencoin, que puede ser usada en blockchain y está pensada para la realización de transacciones internacionales entre bancos de una manera más eficiente (“¿Qué es ripple?”, s.f., párrafos 2-4).

Litecoin: Litecoin es una criptomoneda que fue lanzada el 7 de octubre de 2011 que permite realizar pagos instantáneos y a muy bajo coste a cualquier parte del mundo. Es una red de pagos global completamente descentralizada y de código abierto y se presenta como complemento a Bitcoin (“Litecoin - la moneda electronica”, s.f., párrafo 1).

Monero: Es una criptomoneda descentralizada, creada en abril de 2014, basada en el uso de una blockchain inmutable, con un alto grado de confidencialidad en las transacciones, de tal manera que las transacciones en la blockchain no pueden ser relacionadas con ningún usuario o entidad (“monero - secure, private, untraceable”, s.f., párrafos 2-5)

15.2 Resumen estadístico criptomonedas

Los estadísticos de resumen de los precios diarios y de la rentabilidad continua de las seis criptomonedas están en las tablas 2 y 3 respectivamente. En la Tabla 2 se puede observar como ha variado el valor de las criptomonedas y cuales son las que tienen más valor de mercado. Bitcoin destaca por encima de las demás por tener una media de precio bastante muy superior al del resto de las criptomonedas, además de que el mínimo, el cuartil 1, la mediana, el cuartil 3 y el máximo están por encima de cualquiera de las otras criptomonedas, confirmando que tiene un valor de mercado superior al resto. Por otro lado, el coeficiente de variación indica que Bitcoin es la que tiene menor dispersión relativa de los datos. El resto de criptomonedas tienen valores de al rededor de 0,8, siendo Ripple la que tiene más coeficiente de variación (CV) con un valor de 0.97. Analizando el rango de Bitcoin se puede ver como es aproximadamente tres veces el valor medio, mientras que en el resto de criptomonedas es bastante superior a eso, indicando que han tenido variaciones más importantes en términos relativos respecto al Bitcoin. Por otro lado, todas las criptomonedas tienen una característica en común y es que son asimétricas positivas y leptocúrticas, siendo esto último cierto ya que el coeficiente de curtosis es superior a 0 en todas ellas. Cabe destacar por último que Ripple tiene un apuntamiento muy importante debido a que su coeficiente de Curtosis es de 17.78, muy superior al del resto de criptomonedas.

Estadísticos	Bitcoin	Ethereum	Eth_classic	Litecoin	Monero	Ripple
Media	6337.86	296.87	11.38	75.05	103.72	0.40
Desv. Estándar	3567.40	247.74	8.72	58.86	88.18	0.38
Mínimo	755.88	8.10	1.13	3.72	10.56	0.01
Cuartil 1	3656.20	141.93	4.68	41.61	48.50	0.22
Mediana	6484.45	218.54	8.96	57.99	80.84	0.31
Cuartil 3	8440.51	360.75	16.06	92.50	121.15	0.45
Máximo	19587.37	1452.58	46.78	357.02	509.58	3.33
Asimetría	0.58	1.69	1.30	1.65	1.82	3.57
Curtosis	0.46	3.04	1.54	3.23	3.26	17.78
Rango	18831.49	1444.48	45.65	353.30	499.02	3.32
Varianza	12726354.91	61377.02	76.08	3464.45	7774.87	0.15
CV	0.56	0.83	0.77	0.78	0.85	0.97

Tabla 2: Estadísticos de resumen del precio de cierre de Bitcoin, Ethereum, Ethereum Classic, Ripple, Litecoin y Monero en dólares, del 1 de enero de 2017 al 31 de diciembre de 2019

La tabla 3 da un resumen de los estadísticos principales de la rentabilidad continua de cada una de las criptomonedas. Ripple y Ethereum son las criptomonedas que tienen un valor medio más elevado de la rentabilidad continua mientras que Ethereum Classic y Monero son las que lo tienen más bajo. Ripple destaca por tener el mínimo y el máximo más extremos, concordando con lo visto en la tabla 2. Otro hecho diferenciador de Bitcoin es que es la única criptomoneda que tiene asimetría negativa, mientras que en el caso del apuntamiento, todas las criptomonedas son leptocúrticas y por lo tanto tienen un apuntamiento superior al de la distribución normal. Finalmente, atendiendo a la dispersión relativa, todas destacan por tener una alta dispersión, siendo Ethereum Classic y Monero las que tienen mayor dispersión relativa.

Estadísticos	Bitcoin	Ethereum	Eth_classic	Litecoin	Monero	Ripple
Media	0.0018	0.0025	0.0010	0.0021	0.0011	0.0031
Desv. Estándar	0.0440	0.0584	0.0690	0.0627	0.0633	0.0797
Mínimo	-0.2397	-0.3335	-0.4473	-0.4060	-0.2918	-0.6316
Cuartil 1	-0.0162	-0.0223	-0.0254	-0.0261	-0.0278	-0.0256
Mediana	0.0028	-0.0000	-0.0007	-0.0011	-0.0009	-0.0020
Cuartil 3	0.0198	0.0259	0.0269	0.0264	0.0287	0.0217
Máximo	0.2549	0.2930	0.7022	0.4797	0.4274	1.0198
Asimetría	-0.1390	0.2308	0.7007	0.9665	0.4005	2.7363
Curtosis	4.4792	4.3435	14.3651	9.0648	4.7284	34.7495
Rango	0.4947	0.6265	1.1495	0.8857	0.7192	1.6513
Varianza	0.0019	0.0034	0.0048	0.0039	0.0040	0.0063
CV	24.2009	23.0694	66.5354	30.5534	60.1227	25.8416

Tabla 3: Estadísticos de resumen de la rentabilidad continua de Bitcoin, Ethereum, Ethereum Classic, Ripple, Litecoin y Monero, del 1 de enero de 2017 al 31 de diciembre de 2019

Tal y como se afirma en (Chan et al. 2017), cuando se ajusta una distribución estadística se están haciendo una serie de asunciones, las cuales son que los datos son independientes e idénticamente distribuidos y que no hay presencia de autocorrelación y heteroscedasticidad. Se dan por válidos estos supuestos porque en otros estudios como el mencionado en este párrafo y en (Chu, Nadarajah, and Chan 2015), se concluye que estas hipótesis se sostienen, y de hecho, tal y como se afirma en el segundo, no es nada extraño que la rentabilidad continua de las criptomonedas cumpla estos supuestos.

15.3 Distribuciones de probabilidad y criterios de selección

En este apartado se hace una presentación de las distribuciones de probabilidad que serán ajustadas a la rentabilidad continua de cada criptomoneda, los criterios de información que serán utilizados y el test no paramétrico que servirá para concluir si se puede rechazar o no que los datos provengan de la distribución final seleccionada.

Esto se hace por un lado debido a la asimetría de la mayoría de las distribuciones y a causa del enorme apuntamiento que presentan tal y como se ha visto en la tabla 3. Estas características son opuestas a las de la distribución normal, la cual es mesocúrtica (Curtosis = 0) y simétrica.

Como se ha comentado en el apartado de metodología, este ajuste se hace ya que se tiene como objetivo el cálculo del riesgo asociado a cada criptomoneda (VaR o CVaR) de forma paramétrica, esto es, a partir de los cuantiles de la distribución final seleccionada, por las razones comentadas en dicho apartado. En la figura 21 hay una representación del histograma de las rentabilidades continuas de las seis criptomonedas. En estos histogramas se pueden ver dos cosas: la primera es que ninguno de los histogramas parece indicar, tal y como se ha visto anteriormente con el coeficiente de curtosis, que la rentabilidad continua siga una distribución normal. Por lo tanto, se observa un alto grado de apuntamiento, sobre todo en las criptomonedas Ripple y Ethereum Classic.

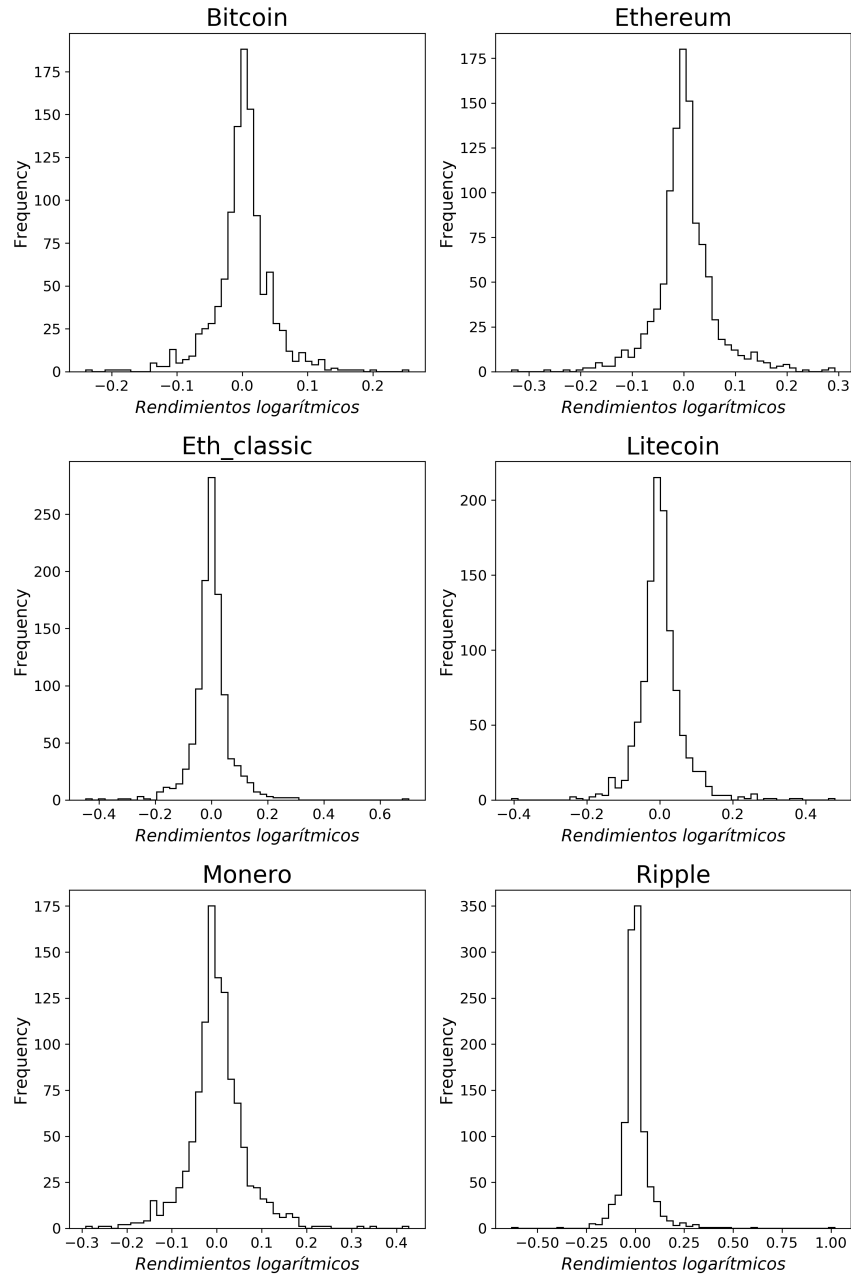


Fig. 21: Histogramas de la rentabilidad continua de las seis criptomonedas del 1 de enero de 2017 al 31 de diciembre de 2019

Si se supone que X denota una variable aleatoria continua que representa la rentabilidad continua de cada una de las criptomonedas. Sea $f(x)$ la función de densidad de probabilidad de X . Sea $F(x)$ la función de distribución de X . A continuación se ajustará a X siete de las distribuciones más usadas en finanzas (Chu, Nadarajah, and Chan 2015), de las que hay disponibles en ecosistema SciPy the Python, a cada una de las criptomonedas.

- Distribución Normal:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{(x - \mu)^2}{2\sigma^2} \right\}.$$

Con $-\infty < x < \infty$, $-\infty < \mu < \infty$ y $\sigma > 0$

- Distribución de Laplace (De Laplace 1774 citado en Chan et al. (2017)):

$$f(x) = \frac{1}{2\sigma} \exp \left\{ -\frac{|x - \mu|}{\sigma} \right\}.$$

Con $-\infty < x < \infty$, $-\infty < \mu < \infty$ y $\sigma > 0$

- La distribución Normal Inversa Gaussiana, en su forma estandarizada (Barndorff-Nielsen 1997 citado en "scipy.stats.norminvgauss", s.f.):

$$f(x, a, b) = (a \exp \sqrt{a^2 - b^2} + bx) / (\pi \sqrt{1 + x^2} K_1(a \sqrt{1 + x^2}))$$

Donde x es un número real, a es la pesadez de cola y b es el parámetro de asimetría, los cuales satisfacen que $a > 0$ y $|b| \leq a$, y donde K_1 es la función modificada de Bessel de segundo tipo. En el propio ecosistema SciPy hay unos parámetros de escala y localización disponibles para ajustar correctamente la distribución a los datos.

- La distribución T de Student (Gosset 1908 citado en Chan et al. (2017)):

$$f(x) = \frac{K(v)}{\sigma} \left\{ 1 + \frac{(x - \mu)^2}{v \sigma^2} \right\}^{-(1+v)/2}$$

Con $-\infty < x < \infty$, $-\infty < \mu < \infty$ y $\sigma > 0$, donde $K(v) = \sqrt{v} B(v/2, 1/2)$ y $B(\cdot)$ es la función beta.

- La distribución Logística (Chu, Nadarajah, and Chan 2015):

$$f(x) = \frac{1}{\sigma} \exp \left\{ -\frac{x - \mu}{\sigma} \right\} \left\{ 1 + \exp \left\{ -\frac{x - \mu}{\sigma} \right\} \right\}$$

Con $-\infty < x < \infty$, $-\infty < \mu < \infty$ y $\sigma > 0$

- La distribución normal sesgada (Azzalini 1985 citado en Chu, Nadarajah, and Chan (2015)):

$$f(x) = \frac{2}{\sigma} \phi \left(\frac{x - \mu}{\sigma} \right) \Phi \left\{ \lambda \frac{x - \mu}{\sigma} \right\}$$

Con $-\infty < x < \infty$, $-\infty < \mu < \infty$ y $\sigma > 0$, donde $\phi(\cdot)$ es la función de distribución de probabilidad y $\Phi(\cdot)$ es la función de distribución acumulativa.

- La distribución exponencial de potencia (Zhu and Zinde-Walsh 2009 citado en Chu, Nadarajah, and Chan (2015))

$$f(x) = \frac{\beta}{2\sigma\Gamma(1/\beta)} \exp \left\{ - \left\{ \frac{|x - \mu|}{\sigma} \right\}^\beta \right\}$$

Con $-\infty < x < \infty$, $-\infty < \mu < \infty$ y $\sigma > 0$, donde $\Gamma(\cdot)$ denota la función gamma. \$

La discriminación entre las distribuciones se ha llevado a cabo con diversos criterios, los cuales son los usados en (Chan et al. 2017):

El criterio de información de Akaike definido por:

$$AIC = 2k - 2\ln L(\hat{\Theta})$$

El criterio de información de Bayes definido por:

$$BIC = k \ln n - 2\ln L(\hat{\Theta})$$

El criterio de información de Akaike consistente definido por:

$$CAIC = -2\ln L(\hat{\Theta}) + k(\ln n + 1)$$

El criterio de información de Akaike corregido definido por:

$$AIC_c = AIC + \frac{2k(k+1)}{n-k-1}$$

El criterio de Hannan-Quinn definido por:

$$HQC = -2\ln L(\hat{\Theta}) + 2k \ln (\ln (n))$$

Cuanto menor sea el valor de cualquier de estos criterios mejor es el ajuste de la distribución de probabilidad. Todos los criterios tienen en común que usan el estimador máximo verosímil. El criterio de AIC está compuesto de dos partes: el sesgo (la log-verosimilitud), que cuanto más grande es, mejor es la bondad del ajuste y la varianza (número de parámetros), que sirve para penalizar el sobreajustamiento. El criterio de BIC tiene como única diferencia que en vez de penalizar con el doble de parámetros del modelo, la penalización es el número de parámetros multiplicado por el logaritmo neperiano del número de individuos.

Continuando con el resto de criterios, y tal y como se indica en (Chan et al. 2017), el criterio de AIC consistente incluye una penalización del sobreajustamiento más estricta que la del criterio de AIC. Por otro lado, el criterio de AIC corregido incluye una corrección que corrige el sesgo en pequeñas muestras y otra penalización que es función del tamaño de la muestra.

Finalmente, el criterio de Hannan-Quinn es parecido al criterio de AIC pero penaliza más la complejidad del modelo, ya que multiplica el producto $2k$ por el doble logaritmo neperiano del tamaño muestral.

Para valorar si los datos provienen de la distribución seleccionada para cada una de las criptomonedas se usa el test de una muestra de Kolmogorov-Smirnov. Tiene como hipótesis nula que los datos provienen de la distribución ajustada y como alternativa que no. Se define un nivel de significación del 5% para rechazar que los datos provienen de la distribución elegida. El estadístico es el siguiente (Prueba de Kolmogorov-Smirnov, s.f.):

$$F_n(X) = \frac{1}{n} \sum_{i=0}^{\infty} \left\{ \begin{array}{l} 1 \text{ si } y_i \leq x, \\ 0 \text{ alternativa.} \end{array} \right\}$$

El test se realiza con la hipótesis alternativa a dos colas. Para dos colas el test viene definido por:

$$D_n^+ = \max(F_n(x) - F(x))$$

$$D_n^- = \max(F_n(x) - F(x))$$

Donde $F(x)$ es la distribución presentada como hipótesis.

15.3.1 Distribuciones seleccionadas y parámetros

Las distribuciones mencionadas han sido ajustadas a cada una de las criptomonedas y se ha escogido la mejor distribución de cada una de ellas atendiendo al criterio de AIC. Los parámetros de estas distribuciones han sido estimados mediante el método de máxima verosimilitud. A continuación, en la Figura 22, se muestra para cada una de las criptomonedas el histograma de sus valores, la función de distribución ajustada y además el valor de los parámetros de cada una de las distribuciones. Como se puede observar, todas las criptomonedas tienen como la distribución que mejor ajusta la normal inversa gaussiana, donde “a” es el peso de la cola, “b” es el parámetro de simetría. Loc y Scale son los parámetros de localización y escala respectivamente.

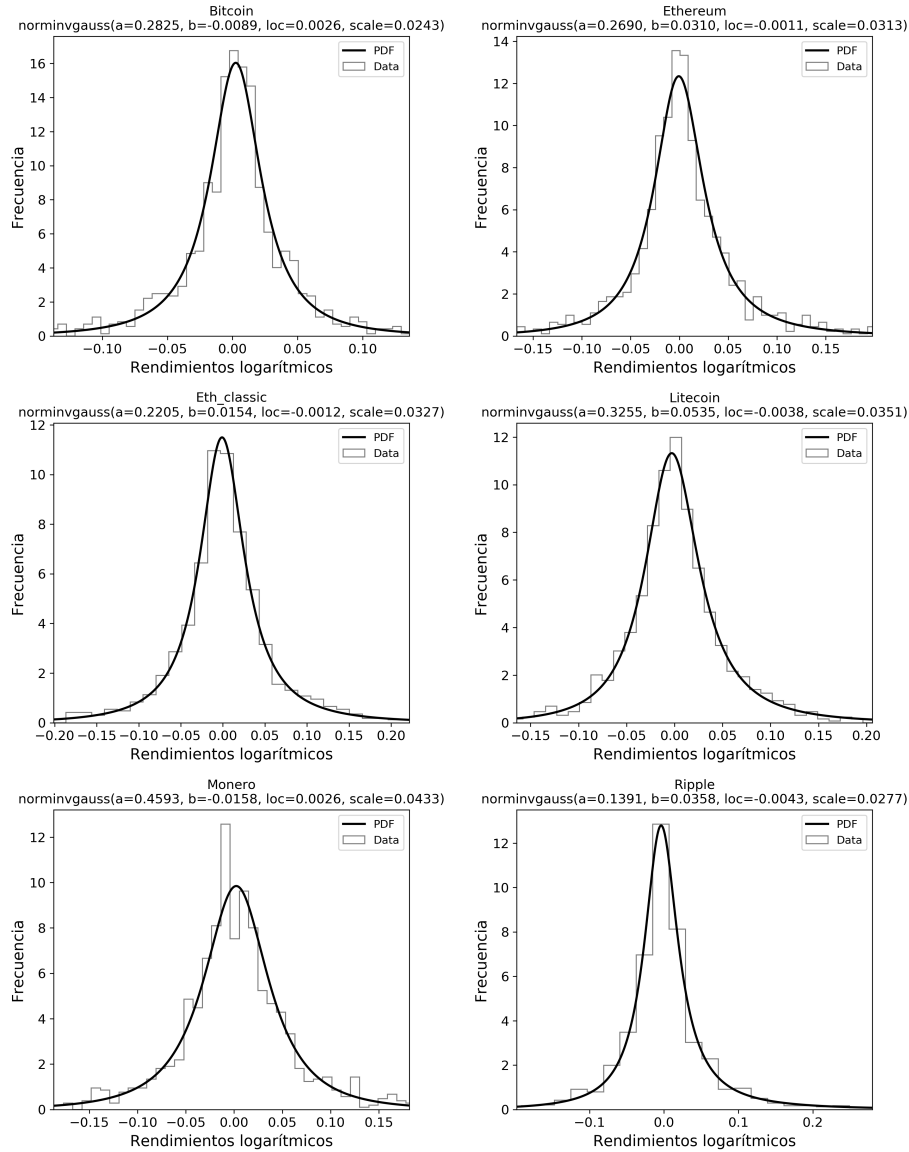


Fig. 22: Histogramas de la rentabilidad continua de las seis criptomonedas junto a la función de distribución de probabilidad que mejor las ajusta y los parámetros de las mismas del 1 de enero de 2017 al 31 de diciembre de 2019

15.3.2 Criterios de información

Se presentan los resultados de los criterios de información para cada una de las criptomonedas en las Tablas 4 a 9. El criterio de AIC ha sido el escogido como prioritario, tal y como se hace en (Chan et al. 2017), pero aún así se muestran los resultados con el resto de criterios. El resto de criterios se muestran por si alguien da prioridad a otros criterios y quiere ver que modelo se hubiera escogido bajo la consideración un criterio distinto al de AIC. Todo esto tiene la utilidad de que, en vez de partir de la suposición de normalidad para posteriormente calcular el VaR paramétrico, se explora cual es la distribución que mejor ajusta los datos. Si en vez de eso se tomara la distribución normal como la distribución de la que provienen los datos, toda conclusión sobre el riesgo asociado a cada

criptomoneda sería incorrecta como se verá a continuación. Esto es debido a que el cálculo del VaR y CVaR estará en función de la distribución final escogida (la que mejor ajuste proporcione). Si ésta no es una distribución de probabilidad que ajusta correctamente la rentabilidad continua entonces los VaRs y CVaRs calculados estarían sesgados, por lo tanto, hará falta un test de bondad de ajuste para asegurarse de que esto no pasa.

Tabla 4: Distribuciones ajustadas y resultados de los criterios de información para el ajuste de rentabilidad continua de Bitcoin del 1 de enero de 2017 al 31 de diciembre de 2019

Distribuciones	- ln L	AIC	AICC	BIC	HQC	CAIC
T Student	-2002.51	-3999.02	-3984.02	-3981.02	-3999.00	-3993.34
Laplace	-2011.05	-4018.11	-4008.11	-4006.11	-4018.10	-4014.32
NormInvGauss	-2013.58	-4019.16	-3999.16	-3995.16	-4019.12	-4011.59
Normal	-1867.99	-3731.98	-3721.98	-3719.98	-3731.97	-3728.20
Logística	-1956.53	-3909.07	-3899.07	-3897.07	-3909.06	-3905.29
Expon. Potencia	-1574.78	-3143.56	-3128.56	-3125.56	-3143.54	-3137.88
NormalSesgada	-1868.14	-3730.28	-3715.29	-3712.29	-3730.26	-3724.61

En la tabla 4 se puede ver como la Normal Inversa Gaussiana es mejor que el resto de las distribuciones según los criterios de AIC y HQC. La distribución de Laplace por otro lado es la mejor según el resto de criterios.

Tabla 5: Distribuciones ajustadas y resultados de los criterios de información para el ajuste de rentabilidad continua de Ripple del 1 de enero de 2017 al 31 de diciembre de 2019

Distribuciones	- ln L	AIC	AICC	BIC	HQC	CAIC
T Student	-1622.79	-3239.57	-3224.58	-3221.58	-3239.55	-3233.90
Laplace	-1571.21	-3138.41	-3128.42	-3126.42	-3138.40	-3134.63
NormInvGauss	-1632.51	-3257.03	-3237.03	-3233.03	-3256.99	-3249.46
Normal	-1217.03	-2430.05	-2420.06	-2418.06	-2430.04	-2426.27
Logística	-1473.41	-2942.83	-2932.83	-2930.83	-2942.81	-2939.04
Expon. Potencia	-372.78	-739.56	-724.57	-721.57	-739.54	-733.89
NormalSesgada	-1277.88	-2549.76	-2534.77	-2531.77	-2549.74	-2544.09

En la tabla 5 se ve claramente como para todos los criterios la Normal Inversa Gaussiana es la distribución que mejor ajusta la rentabilidad continua de Ripple, ya que en todos los criterios tiene el menor valor.

Tabla 6: Distribuciones ajustadas y resultados de los criterios de información para el ajuste de rentabilidad continua de Ethereum del 1 de enero de 2017 al 31 de diciembre de 2019

Distribuciones	- ln L	AIC	AICC	BIC	HQC	CAIC
T Student	-1702.96	-3399.92	-3384.93	-3381.93	-3399.90	-3394.25
Laplace	-1707.43	-3410.86	-3400.86	-3398.86	-3410.85	-3407.07
NormInvGauss	-1715.15	-3422.30	-3402.31	-3398.31	-3422.27	-3414.74
Normal	-1557.55	-3111.09	-3101.10	-3099.10	-3111.08	-3107.31
Logística	-1649.39	-3294.78	-3284.78	-3282.78	-3294.77	-3291.00
Expon. Potencia	-1256.72	-2507.44	-2492.45	-2489.45	-2507.42	-2501.77
NormalSesgada	-1563.69	-3121.39	-3106.39	-3103.39	-3121.36	-3115.71

En la tabla 6 se ve como en el caso de Ethereum en todos los criterios la normal inversa gaussiana es mejor que el resto de distribuciones a excepción de con el criterio de BIC donde la distribución de probabilidad de Laplace tiene un menor valor.

Tabla 7: Distribuciones ajustadas y resultados de los criterios de información para el ajuste de rentabilidad continua de Ethereum Classic del 1 de enero de 2017 al 31 de diciembre de 2019

Distribuciones	- ln L	AIC	AICC	BIC	HQC	CAIC
T Student	-1598.37	-3190.73	-3175.74	-3172.74	-3190.71	-3185.06
Laplace	-1587.42	-3170.85	-3160.85	-3158.85	-3170.83	-3167.06
NormInvGauss	-1606.67	-3205.33	-3185.34	-3181.34	-3205.29	-3197.77
Normal	-1373.64	-2743.28	-2733.29	-2731.29	-2743.27	-2739.50
Logística	-1519.83	-3035.65	-3025.66	-3023.66	-3035.64	-3031.87
Expon. Potencia	-749.78	-1493.55	-1478.55	-1475.55	-1493.53	-1487.88
NormalSesgada	-1383.63	-2761.25	-2746.26	-2743.26	-2761.23	-2755.58

En la tabla 7 se ven los resultados para Ethereum Classic donde la normal inversa gaussiana es la que mejor ajusta según todos los criterios.

Tabla 8: Distribuciones ajustadas y resultados de los criterios de información para el ajuste de rentabilidad continua de Litecoin del 1 de enero de 2017 al 31 de diciembre de 2019

Distribuciones	- ln L	AIC	AICC	BIC	HQC	CAIC
T Student	-1658.90	-3311.80	-3296.80	-3293.80	-3311.78	-3306.13
Laplace	-1652.71	-3301.42	-3291.42	-3289.42	-3301.40	-3297.63
NormInvGauss	-1666.03	-3324.06	-3304.06	-3300.06	-3324.02	-3316.49
Normal	-1478.58	-2953.16	-2943.16	-2941.16	-2953.15	-2949.37
Logística	-1603.32	-3202.65	-3192.65	-3190.65	-3202.63	-3198.86
Expon. Potencia	-977.38	-1948.76	-1933.76	-1930.76	-1948.74	-1943.08
NormalSesgada	-1502.59	-2999.18	-2984.19	-2981.19	-2999.16	-2993.51

En la tabla 8, en la que se ven los resultados de Litecon, se saca la misma conclusión que con Ripple y Ethereum Classic.

Tabla 9: Distribuciones ajustadas y resultados de los criterios de información para el ajuste de rentabilidad continua de Monero del 1 de enero de 2017 al 31 de diciembre de 2019

Distribuciones	- ln L	AIC	AICC	BIC	HQC	CAIC
T Student	-1575.85	-3145.69	-3130.70	-3127.70	-3145.67	-3140.02
Laplace	-1578.37	-3152.73	-3142.73	-3140.73	-3152.72	-3148.95
NormInvGauss	-1580.37	-3152.74	-3132.75	-3128.75	-3152.71	-3145.18
Normal	-1468.40	-2932.80	-2922.80	-2920.80	-2932.79	-2929.02
Logística	-1546.52	-3089.04	-3079.04	-3077.04	-3089.03	-3085.26
Expon. Potencia	-1141.58	-2277.16	-2262.16	-2259.16	-2277.13	-2271.48
NormalSesgada	-1475.74	-2945.48	-2930.49	-2927.49	-2945.46	-2939.81

Finalmente para el caso de Monero, hay discrepancia por cual es la mejor distribución según los criterios que se escojan. A partir del criterio de AIC y AICC se concluye que el mejor modelo es la Normal Inversa Gaussiana. Según el resto de criterios es la distribución de Laplace.

Como se puede observar en todos los casos, la distribución Normal se encuentra entre las distribuciones que, según el criterio de AIC, ajustan peor, lo cual implica que medidas de riesgo habituales como el VaR son imprecisas si se parte del supuesto de normalidad, y esto llevaría seguramente a una infravaloración del riesgo.

15.3.3 Q-Q Plots

Una vez se ha seleccionado que distribución ajusta mejor cada criptomoneda se procede a un análisis de como de correcto es el ajuste.

Los QQ-Plots de la figura 23 muestran como de bueno es el ajuste de las distribuciones que mejor ajustan cada una de las criptomonedas. Estas distribuciones representadas son las escogidas bajo el criterio de AIC a partir de los resultados de las tablas 4-9, que en este caso es la normal inversa gaussiana para las seis criptomonedas.

En el caso de Bitcoin, la distribución ajusta bien en la parte inferior y media de los datos pero no la cola superior. En Ethereum ajusta bien la parte media y la cola superior de los datos pero en la cola inferior tiene problemas para ajustarla. Para Ethereum Classic la distribución ajusta bien el centro de los datos pero la cola inferior no es ajustada correctamente. Se puede observar como hay un dato muy extremo que al ser un outlier no debería considerarse como relevante. Litecoin y Monero tienen un buen ajuste central y de la cola superior pero en la cola inferior hay algunos valores extremos que no ajustan bien. A pesar de ello, la gran mayoría de datos si que son justados como deberían. Finalmente, en el caso de Ripple, el ajuste es bueno en el centro y en la cola inferior.

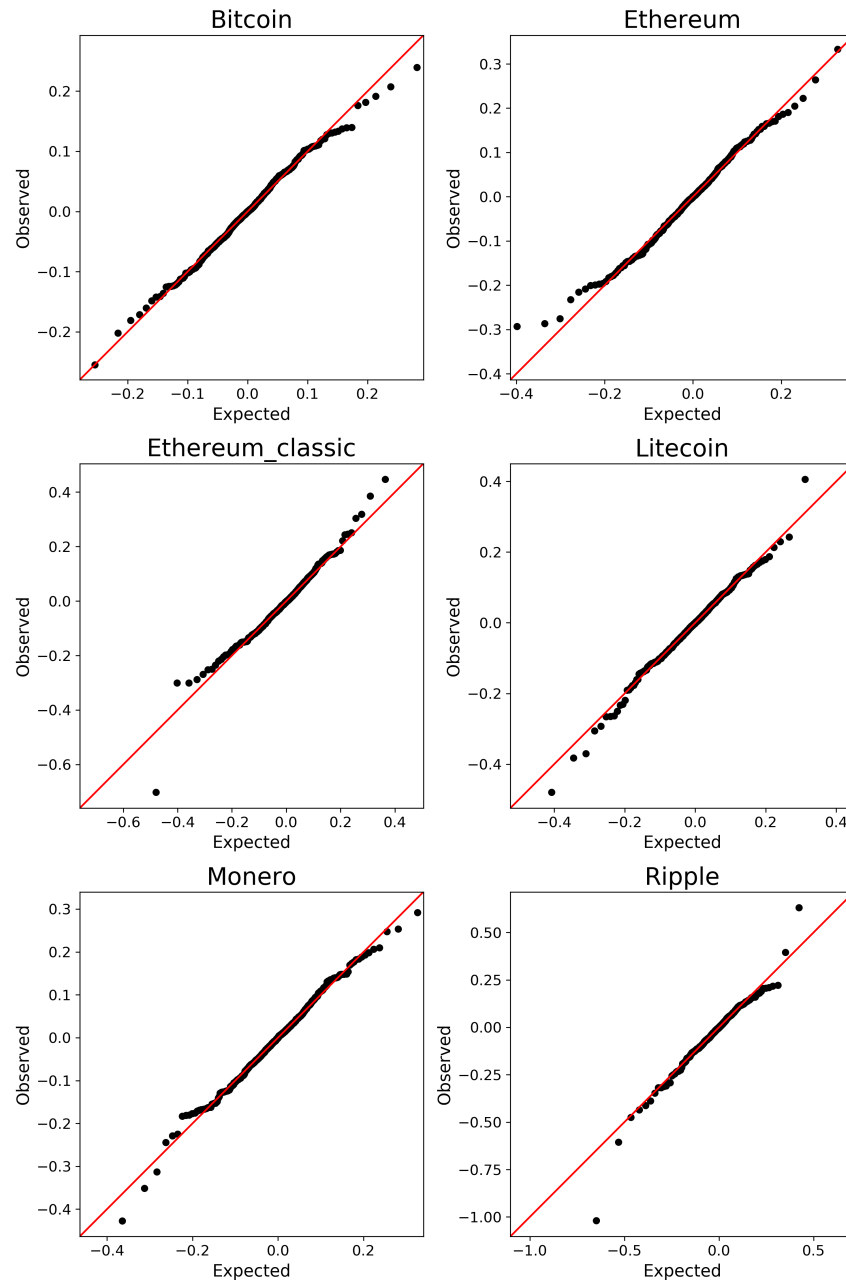


Fig. 23: QQ-Plots de la distribución que mejor ajusta la rentabilidad continua de las seis criptomonedas del 1 de enero de 2017 al 31 de diciembre de 2019

15.3.4 P-P Plots

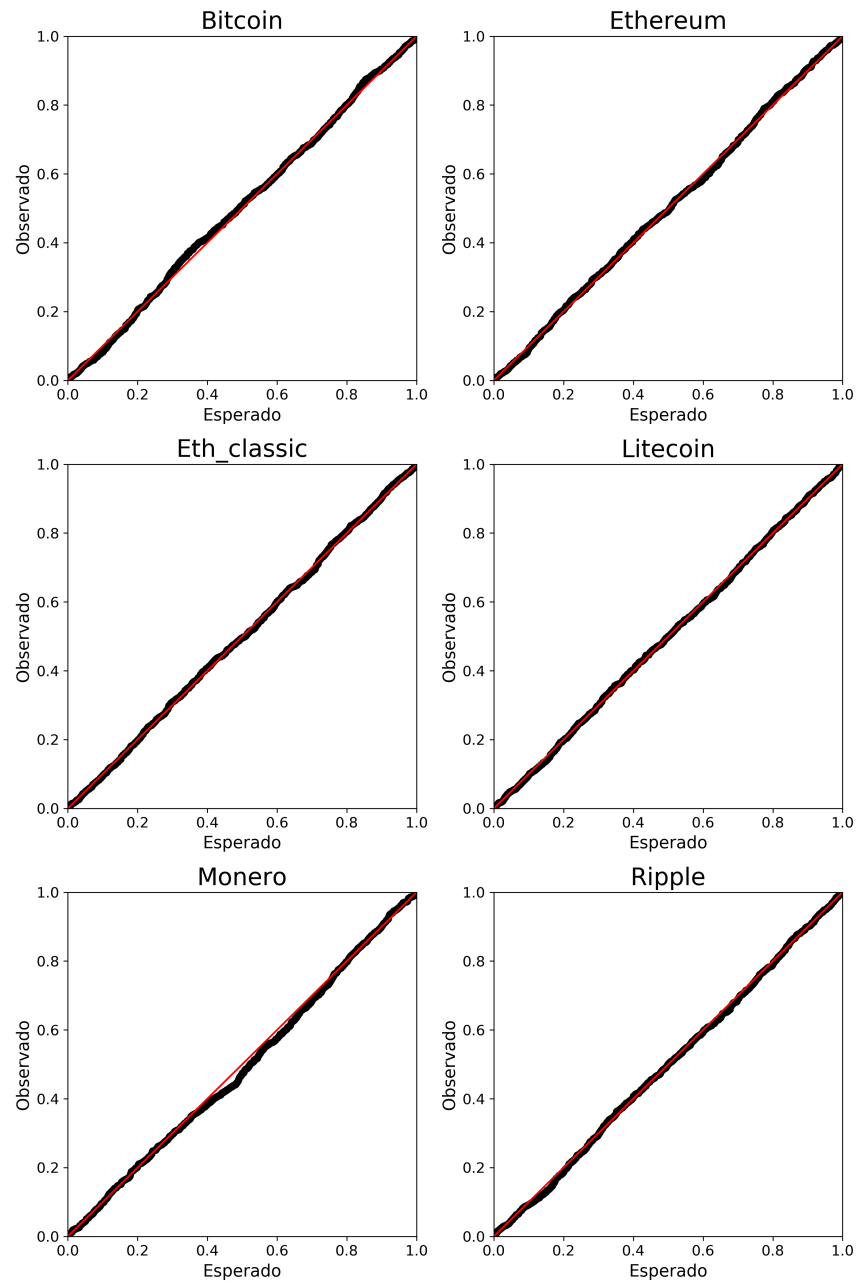


Fig. 24: PP-Plots de la mejor distribución para la rentabilidad continua de las seis criptomonedas del 1 de enero de 2017 al 31 de diciembre de 2019

En la figura 24 se observan los PP-Plots para los datos de la rentabilidad continua de las seis criptomonedas. En todas las criptomonedas la mejor distribución normal invertegaussiana capta la parte central y las dos colas de los datos.

15.3.5 Test Bondad de Ajuste

El test de Kolmogorov Smirnov se ha aplicado a las 6 criptomonedas para la distribución que mejor las ajusta. Los resultados se muestran a continuación en la Tabla 10.

Criptomonedas	KS.P.value
Bitcoin	0.382
Ethereum	0.898
Eth_classic	0.976
Litecoin	0.968
Monero	0.055
Ripple	0.684

Tabla 10: Test de Kolmogorv Smirnov a dos colas para la distribución que mejor ajusta la rentabilidad continua de las 6 criptomonedas del 1 de enero de 2017 al 31 de diciembre de 2019

En la tabla 10 se ve como todos los P-valores son superiores a 0.05 y por lo tanto no se puede rechazar que la rentabilidad continua de cada una de las criptomonedas provenga de la distribución que mejor las ajusta. Pese a ello, cabe destacar que en el caso de Monero no se ha rechazado por muy poco.

15.3.6 VaR plots

El valor en la gestión del riesgo financiero es una medida de riesgo que se utiliza para medir el riesgo de un activo o una cartera y este se define como (Valor en riesgo, s.f.):

Un valor límite tal que la probabilidad de que una pérdida a precios de mercados en la cartera sobre un el horizonte temporal dado exceda ese valor (asumiendo mercados normales y que no se produce negociación en la cartera) sea el nivel de probabilidad dado.

Si $\hat{F}(\cdot)$ representa la función de distribución de la distribución que mejor ajusta cada criptomoneda (según el criterio de AIC y a partir de los resultados de las tablas 4-9), entonces el VaR correspondiente a la probabilidad q puede ser definido de esta forma (Chan et al. 2017):

$$\text{VaR}(q) = \hat{F}^{-1}, \text{ para } 0 < q < 1$$

Ha sido mediante esta fórmula que se han calculado todos los VaR(q).

En la figura 25 se presenta el valor en riesgo de las seis criptomonedas para todos los cuantiles utilizando la distribución que mejor ajusta la rentabilidad continua. Se opta por solo mostrar los de esta distribución, ya que ésta es la que garantizará que las conclusiones sobre el riesgo sean fiables. Se puede ver como para la cola superior, Ripple es la criptomoneda que aparenta tener más riesgo debido a que tiene el VaR más elevado y respecto a la cola inferior, son Ripple, Monero y Ethereum Classic, las cuales tienen un riesgo parecido, las que aparentan tener mayor riesgo. Esto se puede observar ya que para los cuantiles más cercanos a 0 estas últimas tres criptomonedas tienen un VaR(q) inferior al del resto. La criptomoneda con menor riesgo en las dos colas es claramente Bitcoin, que tal y como se había mostrado anteriormente, para los años estudiados en este trabajo, es la criptomoneda con menor coeficiente de variación, aspecto que encaja con el

mostrado en esta figura. Realmente el interés está en los $\text{VaR}(q)$ negativos ya que son estos los que indican mayor riesgo a sufrir grandes pérdidas.

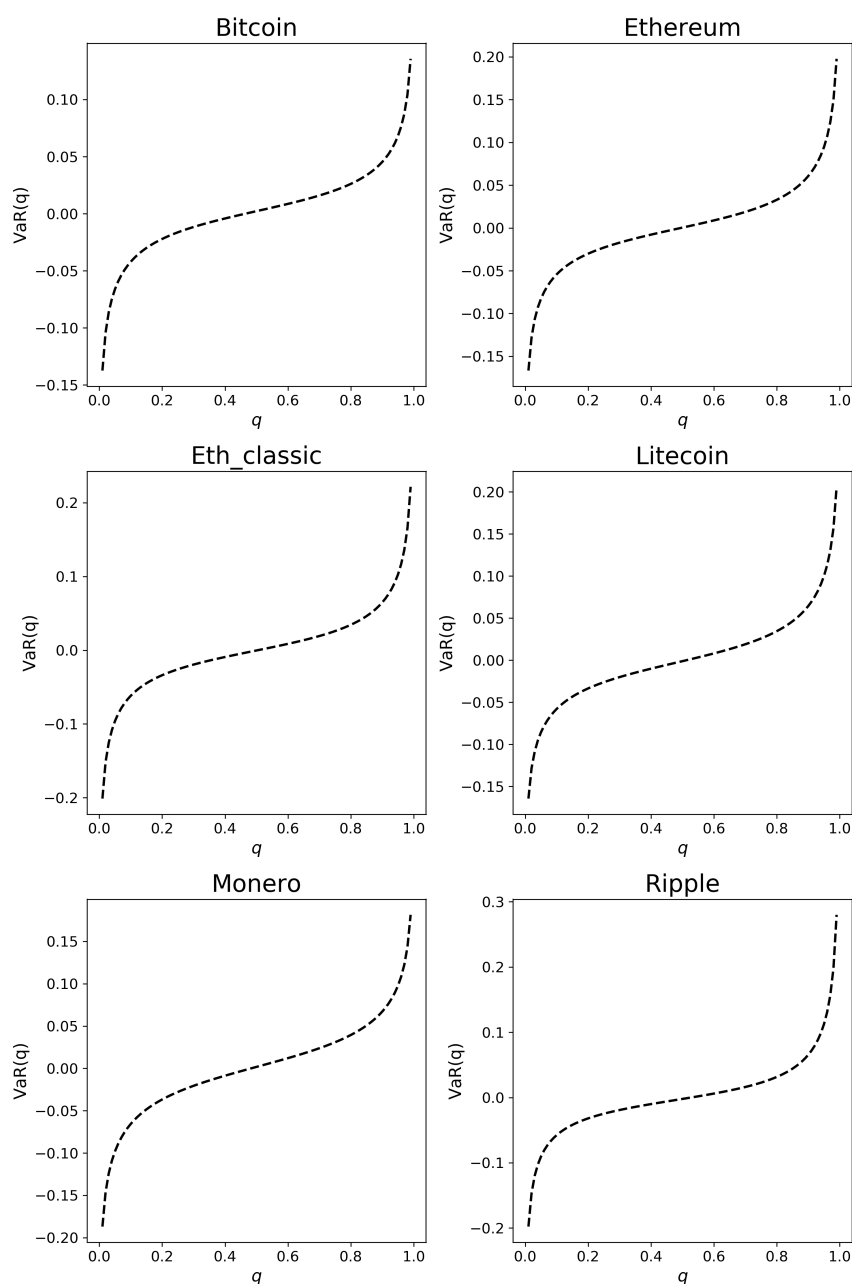


Fig. 25: VaR para la distribución que mejor ajusta los datos de la rentabilidad continua de las seis criptomonedas del 1 de enero de 2017 al 31 de diciembre de 2019

15.3.7 CVaR plots

El CVaR, medida adecuada para variables con “fat tails”, mide la rentabilidad continua esperada cuando la rentabilidad continua es inferior al VaR para una q determinada. Definido

formalmente, es lo siguiente (Chan et al. 2017):

$$ES(q) = \frac{1}{q} \int_0^q VaR(u) du$$

, para $0 < q < 1$

Debido a la enorme dificultad de resolver esta integral para cada q y dada la falta de herramientas para calcular el CVaR para diferentes cuantiles de la normal inversa gaussiana, se opta por la metodología de bootstrap paramétrico.

Se han hecho un millón de simulaciones de valores generados por la función de probabilidad $f_{\hat{\theta}}$ que mejor ajusta cada criptomoneda donde $\hat{\theta}$ son los parámetros máximo verosímiles. A partir de estas simulaciones, y con los VaRs que se habían calculado en el apartado anterior de forma exacta, se han calculado los CVaR(q) para $q \in [0.01, 0.02, \dots, 1]$. Se podrían haber recalculado los VaR(q) a partir de estos datos generados, y con estos VaR(q) calcular cada CVaR(q), pero debido a que se han simulado un millón de valores para cada criptomoneda, los resultados que se hubieran obtenido hubiesen sido prácticamente los mismos por lo que se opta por reusarlos.

En la figura 26 se presenta el CVaR de las seis criptomonedas para todos los cuantiles utilizando la distribución que mejor ajusta la rentabilidad continua. Ethereum Classic y Ripple son las criptomonedas que aparentan tener más riesgo ya que sus CVaR(q) son más extremos para q próximas a 0, siendo Ripple la criptomoneda que tiene más. Esto se puede observar ya que para las cuantiles próximos a 0 sus CVaR(q) son bastante inferiores al del resto de criptomonedas.

Por otro lado, Bitcoin es la que tiene menor riesgo debido a que, para los cuantiles inferiores, tiene unos CVaR(q) claramente inferiores en comparación con el resto. Por último, Ethereum, Monero y Litecoin se encuentran en un punto medio, aparentando no tener tanto riesgo como las dos primeras criptomonedas comentadas, ni tan poco como Bitcoin.

Cabe mencionar que este riesgo ha estado medido a partir de un periodo concreto, que es el comprendido entre el 1 de enero de 2017 y el 30 de diciembre de 2019, por lo tanto, aunque estos gráficos son un buen indicador de la situación actual, no hay ninguna garantía de que esta situación no cambie en los próximos años y que por lo tanto, estas conclusiones queden invalidadas.

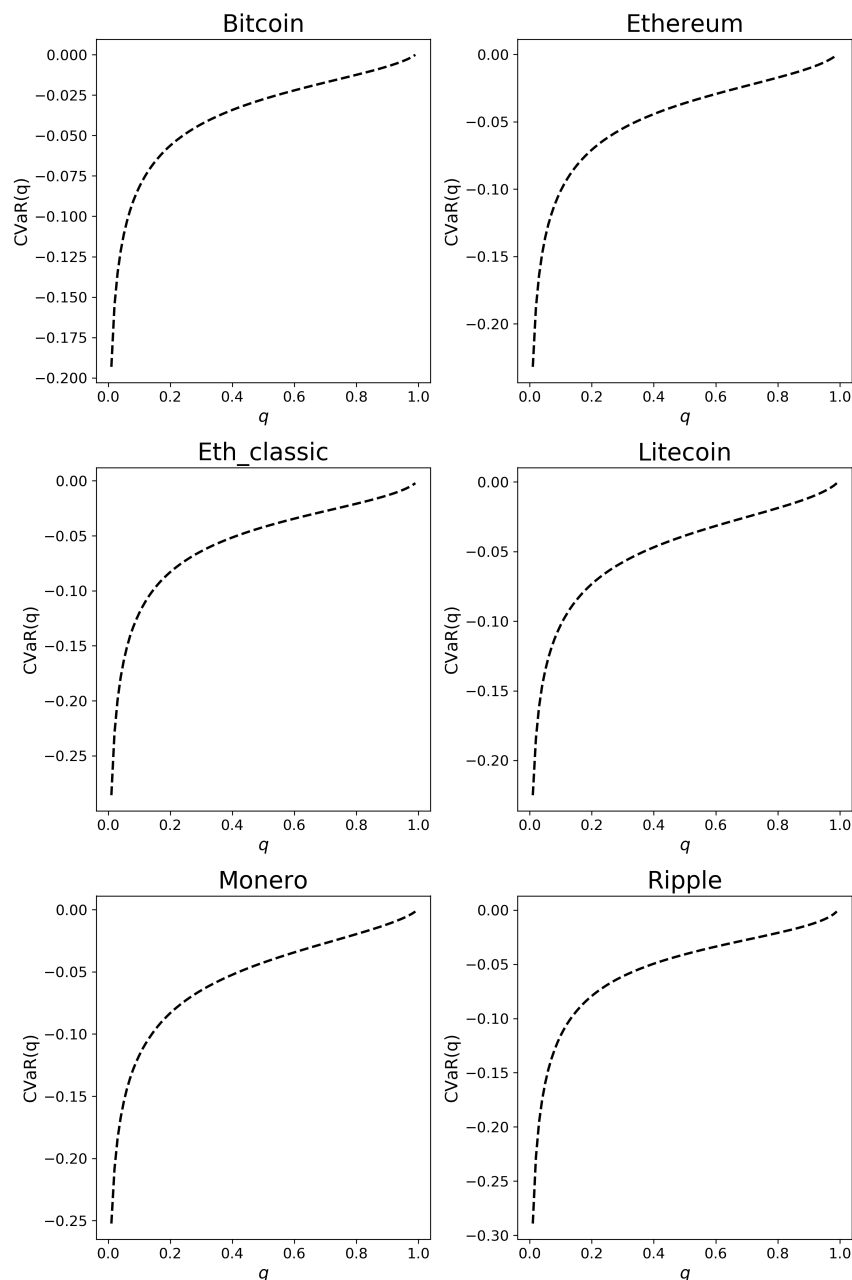


Fig. 26: CVaR para la distribución que mejor ajusta los datos de la rentabilidad continua de cada una de las seis criptomonedas del 1 de enero de 2017 al 31 de diciembre de 2019

Una vez analizado el riesgo asociado a cada criptomoneda a través del VaR y CVaR, la siguiente parte del trabajo está dedicada al ajuste de modelos predictivos para tratar de predecir el precio de cierre de las criptomonedas. Los modelos utilizados serán el modelo ARIMA, que actúa como *benchmark*, y las redes LSTM. A partir de las predicciones realizadas por dicho modelos se podrá establecer que modelo tiene mayor capacidad de predicción de los precios futuros. Será con estas predicciones que se estudiarán los beneficios que pueden extraerse mediante una estrategia de trading sencilla.

Una vez hecho todo esto se dispondrá para cada criptomoneda de información de por un lado, los riesgos asociados a las mismas, obtenidos a partir del análisis estadístico que se acaba de realizar y por otro, de los beneficios resultantes de la aplicación de la estrategia de trading. Con esto se consigue tener un análisis del beneficio-riesgo de cada criptomoneda.

16 Algoritmos de ML vs modelos estadísticos

Tal y como se menciona en (Athey and Imbens 2019), los algoritmos de ML han visto incrementado su uso en problemas en los que uno se enfrenta a datos muy complejos y estos han sido aceptados cada vez más por la comunidad estadística, pero en economía la aceptación está siendo lenta. Esto es debido a que las revistas económicas ponen énfasis en propiedades de los estimadores como son la consistencia, la normalidad y la eficiencia. Por otro lado, los algoritmos de ML ponen más el foco en crear modelos que minimicen una medida de error en datos que no han sido ajustados. Hay situaciones en las que con el análisis de la bondad de ajuste de un modelo es suficiente y otras donde es necesario tener una medida de confianza como los intervalos de confianza. Por otro lado, también existen situaciones en las que lo que se necesita es tener un modelo con una alta capacidad predictiva, que no suele ser reconocido explícitamente como el objetivo principal en la econometría tal y como se afirma en (Athey and Imbens 2019), y es en estas situaciones en las que los algoritmos de ML tienen cabida como una alternativa a los modelos estadísticos.

Mientras que los estadísticos y econométricos ponen el foco en la relación causal entre las covariables y la variable dependiente, la endogeneidad de las variables o la esfericidad del término de perturbación, en el caso de los algoritmos de Machine Learning estos conceptos no son tomados en consideración. En vez de eso, cuando se utiliza un algoritmo de ML para ajustar unos datos, el principal centro de atención es encontrar unos hiperparámetros (parámetros cuyo valor se establece antes de que comience el proceso de aprendizaje), los cuales garantizan tener un modelo que sea capaz de generalizar y hacer buenas predicciones en datos que todavía no ha visto.

La diferencia entre la manera de calcular los parámetros se puede expresar formalmente. En estadística, si se modeliza la distribución condicionada de una variable endógena Y_i dado un regresor X_i y bajo la suposición de que la distribución de la variable endógena condicionada al regresor sigue una distribución normal, esto puede expresarse de la siguiente manera:

$$Y_i|X_i \sim N(\alpha + \beta^T X_i, \sigma^2)$$

Entonces el objetivo es estimar los parámetros $\theta = (\alpha, \beta)$ mediante el método de los mínimos cuadrados ordinarios, que formalmente puede ser expresado de la siguiente manera:

$$(\hat{\alpha}_{mq}, \hat{\beta}_{mq}) = \arg \min_{\alpha, \beta} \sum_{j=1}^{p-1} (Y_i - \alpha - \beta^T X_i)^2$$

Este estimador de mínimos cuadrados ordinarios, siempre que se parta de supuestos como esfericidad del término de perturbación o la ausencia de multicolinealidad, tiene cuatro

propiedades:

- **Linealidad:** los estimadores son una combinación lineal de los verdaderos parámetros poblacionales β , las variables explicativas y el término de perturbación.
- **Absencia de sesgo:** su valor esperado coincide con el parámetro poblacional.
- **Eficiencia:** los estimadores MCO son los que tienen variancia mínima.
- **Consistencia:** al incrementar la medida de la muestra los estimadores se acercan a su valor poblacional.

Por otro lado, tal y como se menciona en (Athey and Imbens 2019), en ML el objetivo es minimizar una función, la cual suele ser también el error cuadrático, pero para nuevas observaciones Y_{N+i} para $i \in \{1, 2, \dots, m\}$ para las que conocemos los valores X_{N+i} . Entonces el objetivo es encontrar unos parámetros $((\hat{\alpha}, \hat{\beta}))$ que tengan buenas propiedades asociadas con la función de pérdida escogida.

En (Shmueli and others 2010) se da una discusión sobre el uso de modelos para predecir o explicar y se dan varias aclaraciones como que en los modelos explicativos la prioridad es maximizar la capacidad explicativa que suele calcularse mediante el R^2 y el autor menciona casos de estudios en los que se ha reportado este estadístico como medida de la capacidad predictiva, dándole por lo tanto un uso totalmente erróneo. Además se demuestra como dado dos “constructs” X e Y , que son “fenómenos con un interés teórico” (Edwards and Bagozzi 2000), los cuales pueden ser observables y no observables, una función F que es la función verdadera que los relaciona y una función f que es una operacionalización correcta de F , se puede demostrar que si f^* es una función sesgada de f , f puede ser preferible desde un punto de vista predictivo a f^* , el cual es mejor desde un punto de vista teórico-explicativo.

Para entender estos conceptos se pone el siguiente ejemplo en el que los tres constructs son observables: el construct Y es la altura de una especie animal, X es el peso y r es la esperanza de vida. Suponiendo que F es que la altura de este animal es dos veces su peso más tres veces su esperanza de vida (lo cual se supone que es la relación que se ha establecido que existe) entonces una correcta operacionalización de F es $f = Y = 2X + 3r$.

En (Shmueli and others 2010) para ilustrar esto se pone el ejemplo del modelo estadístico f que se asume que está correctamente especificado respecto a F : $f(x) = \beta_1 x_1 + \beta_2 x_2 + \epsilon$, $\epsilon \sim \text{iid } (0, \sigma_\epsilon^2)$, el cual tiene las siguientes propiedades:

- Sesgo=0
- $\text{Var}(f(\hat{x})) = \text{Var}(x_1 \hat{\beta}_1 + x_2 \hat{\beta}_2) = \text{Var}(x\hat{\beta}) = x' \text{Var}(\hat{\beta})x = \sigma^2 x' (X'X)^{-1}x$

Donde x es el vector $x = [x_1, x_2]'$ y X es la matriz de diseño basada en los dos predictores.

Se parte del supuesto de que la beta se estima a partir de mínimos cuadrados ordinarios, por lo que $\text{var}(\hat{\beta}) = \sigma^2 (X'X)^{-1}$

- Error predictivo esperado = $E(Y - \hat{f}(x))^2 = E[(f(x) + \epsilon - \hat{f}(x))^2] =$
 $= E[(f(x) - \hat{f}(x)) + \epsilon]^2 = E[(\hat{f}(x) - f(x))^2] + 2E[(\hat{f}(x) - f(x))\epsilon] + E[\epsilon^2] =$

$$\begin{aligned}
&= E[(\hat{f}(x) - f(x))^2] + 2(E((\hat{f}(x))^2 - f(x))) \cdot E[\epsilon] + \sigma^2 = E[(\hat{f}(x) - f(x))^2] + \sigma^2 = \\
&= ECM[\hat{f}(x)] + \sigma^2 = Var(\hat{f}(x)) + Sesgo(\hat{f}(x))^2 + \sigma^2 = \sigma^2 x' (X' X)^{-1} x + 0 + \sigma^2 = \\
&= \sigma^2 (1 + x' (X' X)^{-1} x)
\end{aligned}$$

Y considerando el modelo: $\hat{f}^*(x) = \hat{\gamma}_1 x_1$ que se asume que está sesgado y subespecificado, y que por lo tanto no es una correcta operacionalización de F. En esta situación, el sesgo y varianza están calculados en (Montgomery, Peck, and Vining 2001 citado en Shmueli and others (2010)) y son los siguientes:

- Sesgo = $x_1 \gamma_1 - (x_1 \beta_1 + x_2 \beta_2) = x_1(x_1' x_1)^{-1} x_1' (x_1 \beta_1 + x_2 \beta_2) - (x_1 \beta_1 + x_2 \beta_2)$
- $Var(\hat{f}^*) = x_1 Var(\hat{\gamma}_1) x_1 = \sigma^2 x_1 (x_1' x_1)^{-1} x_1$.

Mediante la combinación del sesgo y la varianza el error predictivo esperado es el siguiente:

$$EPE = (x_1 (x_1' x_1)^{-1} x_1' x_2 \beta_2 - x_2 \beta_2)^2 + \sigma^2 (1 + x_1' (X' X)^{-1} x_1)$$

Comparando los dos errores predictivos esperados de ambos modelos se puede observar como el sesgo en el caso del modelo \hat{f}^* es mayor que el de $\hat{f}(x)$, pero la varianza del modelo sesgado puede llegar a ser menor. Es por esto que el error predictivo esperado puede ser menor en el modelo subespecificado en una serie de situaciones descritas en (Wu, Harris, and McAuley 2007 citado en Shmueli and others (2010)):

- los datos tienen mucho ruido (σ grande)
- el valor real absoluto de los parámetros omitidos, en este caso β_2 , es pequeño.
- cuando los predictores están muy correlacionados
- cuando el tamaño muestral es pequeño o el rango de las variables omitidas es pequeño.

Esto demuestra que un modelo que no es una correcta operacionalización de F y que además está subespecificado, condición necesaria para que todo lo dicho aplique, puede tener mejor capacidad predictiva que otro que representa perfectamente la operacionalización de F.

Es debido a esto, que en estadística hay varios enfoques para tratar de encontrar un modelo que no sobreajuste los datos y esto se hace mediante la inclusión de penalizaciones a la complejidad. Ejemplos de esto son el criterio de AIC, que permite que haya un *trade-off* entre la bondad de ajuste del modelo y la complejidad del modelo.

En el caso de ML hay otras metodologías como la regularización. Esta consiste en tratar de encontrar los parámetros que minimizan el error cuadrático pero añadiendo además un hiperparámetro λ que multiplica el sumatorio de los parámetros del modelo, el cual evita que se produzca un sobreajuste.

17 Introducción a la predicción del precio de las criptomonedas

En esta segunda parte del análisis estadístico de las criptomonedas se ajustarán diversos modelos predictivos a cada una de ellas con la finalidad de determinar como de bien son capaces de predecir el precio de cierre futuro. Se evaluarán la raíz del error cuadrático medio de las predicciones del precio de cierre para una ventana temporal de tres meses. Estos datos con los que se evaluará la capacidad predictiva forman parte de lo que se llama como *test set*, los cuales son unos datos que están aparte del resto de datos con los que se entrenan los modelos. De esta forma se consigue no tener unos resultados muy optimistas, ya que los datos en los que se evalúan los modelos nunca han sido utilizados para el ajuste.

Los modelos evaluados, que serán explicados en el siguiente apartado, son: el modelo autorregresivo integrado de promedio móvil, también llamada ARIMA, que se trata de un modelo estadístico cuyas estimaciones futuras se basan en datos del pasado. Por otro lado, se estudiarán las redes LSTM, las cuales están compuestas de unas celdas LSTM que son capaces de preservar información de momentos del tiempo pasados y por lo tanto, a diferencia de las redes neuronales recurrentes estándar, estas tienen la capacidad de aprender dependencias más largas. Una de las diferencias más importantes entre estos dos modelos es que mientras en el primero el output depende linealmente del input (es un modelo lineal), en el segundo el output depende no linealmente del input (es un modelo no lineal).

Por lo tanto, con esta parte del trabajo se podrá concluir si la red LSTM predice mejor o peor que el modelo ARIMA y en caso de que lo haga si es algo generalizado en todas las criptomonedas o solo en algunas.

Ambos métodos solo utilizarán el precio de cierre como única variable y de esta forma se podrá comparar la capacidad predictiva de ambos modelos de una forma “justa”. Cabe destacar que se hubiera querido hacer una comparación ajustando la red LSTM con más variables para analizar el efecto que esto podría tener, pero la capacidad computacional del equipo informático utilizado no lo ha permitido. Estas variables podrían haber sido el precio mínimo, máximo y medio diario (tomando en consideración todas las transacciones realizadas) o el número de transacciones diarias. Todas las variables incluidas tendrían que haber cumplido con el requerimiento de tener una alta correlación con el precio de cierre, ya que en caso contrario, la tarea de aprendizaje se podría haber visto afectada gravemente.

Otro aspecto a comentar es el hecho de que no se haya estudiado el uso de un modelo SARIMA. Se ha descartado su uso debido a que se parte de la hipótesis, más que razonable, que las criptomonedas no tienen ningún tipo de comportamiento estacional sino que se rigen por un comportamiento totalmente independiente de la temporada del año.

Hay varios artículos académicos relativos al ajuste de modelos predictivos a criptomonedas: (Ji, Kim, and Im 2019) donde se hace un estudio de la capacidad predictiva del precio de cierre de Bitcoin con varios modelos de regresión (predecir el precio del día siguiente) y clasificación (predecir si el precio va a subir o bajar) con datos del 29 de noviembre de 2011 al 31 de diciembre de 2018 y con 18 predictores en total. En este se analizan diferentes tamaños de secuencia de entrada m , que en el caso de las series temporales, es el número de periodos previos que se usan para hacer la predicción del precio, en este caso días. Concluyen, a partir de los algoritmos de

ML que usan entre los que se encuentran las redes LSTM, que con una m mayor o igual a 20 días, la predicción se deteriora. En el caso de clasificación, pese que para una m pequeña la red LSTM tiene una exactitud mejor que la mayoría de modelos, para m grande, el resto de modelos obtienen mejores resultados. Estos resultados, medidos con la variable *accuracy* (porcentaje de predicciones de subidas y bajadas de precios correctas), están muy cerca del 51% en todos los modelos. Finalmente cabe destacar el apartado en el que se evalúan los beneficios de diferentes estrategias de compraventa sin considerar tasas. Mediante regresión, todos los modelos, con una inversión inicial de 10000 dólares, tienen pérdidas de hasta más de 4000 dólares. Por otro lado, cuando los modelos son de clasificación, casi todos los modelos tienen beneficios que van de los 359 dólares a los 877 dólares. Concretamente la red LSTM tiene unos beneficios de 359.2 dólares.

En (Struga and Qirici 2018), se describen las características deflacionarias del Bitcoin debido a que su oferta sigue una serie geométrica y la creación de Bitcoins se hace a una tasa constante. Recogieron datos de la blockchain de Bitcoin, una variable del interés del Bitcoin a través del tiempo y variables del S&P y Dow and Jones y ajustaron una red LSTM. Debido a que no hay una distribución de referencia no se pueden sacar conclusiones de como de bueno es el ajuste.

Finalmente, en (Valencia, Gómez-Espinosa, and Valdés-Aguirre 2019) se hace un análisis de la capacidad predictiva de diferentes modelos de Machine Learning para diferentes criptomonedas: Bitcoin, Ethereum, Ripple y Litecoin. En este artículo utilizan cientos de miles de tuits de Twitter recogidos durante 60 días, a partir de los cuales se crean una serie de variables que evalúan los sentimientos asociados a cada tuit. La conclusión que se saca es que es posible predecir el precio de las criptomonedas a través de Twitter pero con los datos del mercado se obtienen resultados parecidos a los obtenidos usando datos de Twitter y el Mercado conjuntamente.

17.1 Preparación antes del ajuste de los modelos

17.1.1 Redes LSTM

Se ha utilizado el precio de cierre de todas las criptomonedas del 1 de enero de 2017 al 31 de diciembre de 2019, estos son 1095 días.

A continuación se define la notación que se usará para las redes LSTM bajo el supuesto, que es el que se da en este trabajo, de que solo se usa el precio de cierre.

- **R**: es un vector con los precios de cierre original.
- **R[i:j]**: es la secuencia de los precios que van de R[i] a R[j]
- **S**: es una tabla que contiene 1095 - m filas, donde cada fila es una secuencia de tamaño $m+1$. Los elementos de cada fila se expresan de la siguiente manera: $\{s_1, s_2, \dots, s_m, s_{m+1}\}$. El elemento s_{m+1} es el elemento de cada fila que se quiere predecir a partir de los elementos que van de s_1 a s_m . Hay 1095 - m filas, ya que las m primeras filas se eliminan al contener missings. Esto es debido a que se ha limitado la ventana temporal al periodo definido al inicio de este apartado y si se quiere hacer por ejemplo una predicción para el precio del 2 de enero a partir de los cinco precios anteriores, no se dispone de los precios de los días 29, 30 y 31 de diciembre de 2016. Esto se ha hecho para que haya homogeneidad respecto al número de datos que se utilizan en el modelo ARIMA y el resto.

- $S[t,v]$: elemento v de la secuencia número t de la tabla S donde t va de 1 hasta $(1095-m)$ y v va de 1 hasta $m+1$.

Para la preparación de los datos se han seguido dos métodos. En el caso de los algoritmos de Machine Learning, como son las redes LSTM, el procedimiento ha sido el siguiente:

Primero de todo se trabaja a partir del precio de cierre de cada una de las criptomonedas. En la tabla 11 se pueden observar los precios de cierre de Bitcoin del 02/01/2017 al 08/01/2017, rango que se ha escogido al azar para así servir como ejemplo. Estos precios son transformados para que el problema a resolver se convierta en uno supervisado tal y como muestra la tabla 12.

Fecha	Close
2017-01-02	1059.71
2017-01-03	1067.07
2017-01-04	1229.90
2017-01-05	1025.87
2017-01-06	902.93
2017-01-07	914.54
2017-01-08	918.37

Tabla 11: Precios de cierre de Bitcoin para el periodo comprendido entre el 2 de enero de 2017 y el 8 de enero de 2017.

En la tabla 12, que es una muestra de la tabla S para $m=3$, lo que se ha hecho es crear secuencias con una $m=3$, que significa que para predecir el precio para cada una de las fechas señaladas en la primera columna, esto es, el precio de la columna Precio.t, se utilizan los precios de los tres días anteriores. Por ejemplo, para predecir el precio del 4 de enero de 2017 que es S_{m+1} , el cual es de 1229.90, se utilizarán los precios de los días que van del 1 de enero al 3 de enero de 2017.

Fecha	Precio.t.3.	Precio.t.2.	Precio.t.1.	Precio.t.
2017-01-04	1030.65	1059.71	1067.07	1229.90
2017-01-05	1059.71	1067.07	1229.90	1025.87
2017-01-06	1067.07	1229.90	1025.87	902.93
2017-01-07	1229.90	1025.87	902.93	914.54
2017-01-08	1025.87	902.93	914.54	918.37

Tabla 12: Precios de cierre de la tabla 11 en formato supervisado.

Una vez se ha convertido el problema a supervisado mediante la creación de estas secuencias entonces se tiene que llevar a cabo la normalización. Esta es absolutamente necesaria, sobre todo para criptomonedas como Bitcoin, ya que, tal y como se observó en la tabla 1, el precio mínimo fue de 755 dólares y el máximo de 19587 dólares. Una opción es hacer la normalización *minmax*, presente en la fórmula 1. Esta consiste en aplicar para todo elemento i del vector R la siguiente fórmula:

$$R_{MinMax}[i] = \frac{R[i] - R_{min}}{R_{max} - R_{min}}(1)$$

Tal y como se apunta en (Ji, Kim, and Im 2019), si se lleva a cabo la “minmax normalization”, descrita en la fórmula 1, se consigue que todos los datos estén entre 0 y 1. Pero lo que sucede es que cuando la/s variable/s que se normalizan varían mucho, como es este caso, los precios normalizados con este método prácticamente no cambian. Es por esto que se llevará a cabo una normalización por secuencia, descrita en la fórmula 2. Esta normalización consiste en para cada fila t de la tabla S , aplicar para todo elemento v la operación descrita en la fórmula 2 que consiste en calcular en cada secuencia las diferencias de precio respecto al precio inicial de cada una de ellas.

$$S_{norm}[t, v] = \frac{S[t, v]}{S[t, 0]} - 1(2)$$

Por lo tanto, con la normalización minmax lo que se predecirá no es el precio de cierre del elemento $m+1$ de cada fila t de la tabla S , sino la variación de este precio respecto al primer elemento de dicha fila. Un ejemplo de esto se puede ver en la tabla 13, en la que se ha aplicado una normalización minmax de la tabla 12. Este método es mejor que el anterior mencionado porque tal y como comentan en (Ji, Kim, and Im 2019) es capaz de captar mejor las tendencias locales.

Fecha	Precio.t.3.	Precio.t.2.	Precio.t.1.	Precio.t.
2017-01-04	0.0000	0.0282	0.0353	0.1933
2017-01-05	0.0000	0.0069	0.1606	-0.0319
2017-01-06	0.0000	0.1526	-0.0386	-0.1538
2017-01-07	0.0000	-0.1659	-0.2659	-0.2564
2017-01-08	0.0000	-0.1198	-0.1085	-0.1048

Tabla 13: Precios supervisados de la tabla 12 a los que se les ha aplicado la normalización de la fórmula 2.

17.1.2 Definición de las particiones

La tabla S se ha dividido en tres particiones que son:

- *reduced training set*, que contiene los datos que van del 01/01/2017 al 19/10/2018 (60% de los datos).
- *validation set*, que contiene los datos que van del 20/10/2018 al 26/05/2019 (20% de los datos)
- *test set*, que contiene los datos que van del 27/05/2019 al 31/12/2019 (20% de los datos).

Los datos del training set serán utilizados para ajustar diferentes modelos y el validation set para evaluar como de bien predicen. Una vez se ha encontrado, para cada criptomoneda, el modelo que mejor ajusta en el validation set entonces se ajustarán los datos del *training set* (reduced training set + validation set) con el mejor modelo seleccionado y se evaluará la capacidad predictiva final en el test set, datos que aún no habrán sido utilizados hasta ese momento.

Se ha visto en otros artículos académicos como se encontraba el modelo que mejor ajustaba el conjunto de los datos para después reportar la capacidad predictiva del mismo. El problema que esto supone es que se acaba escogiendo un modelo que ha sido entrenado con los datos en los que evalúa su capacidad predictiva y por lo tanto los modelos ajustados no generalizan bien en nuevos datos. La garantía de que se tiene un modelo que generalice bien en nuevos datos es siempre evaluar su capacidad predictiva en datos que aún no ha visto, porque en caso contrario, puede haberse producido lo que se llama un sobreajustamiento de los datos; tienes un modelo que para los datos con los que ha sido entrenado predice a la perfección, pero en cuanto recibe datos con los que nunca ha entrenado las predicciones son muy pobres.

17.1.3 Modelo ARIMA

En el caso del modelo ARIMA, la forma de proceder ha sido distinta a la explicada en los anteriores párrafos. El modelo ARIMA tiene la particularidad de que requiere, tal y como se comenta en (Azari 2019), donde utilizan el modelo ARIMA para llevar a cabo predicciones del precio de Bitcoin, que el conjunto de datos sea estacionario. Esta estacionalidad se ha conseguido llevando a cabo diferenciaciones de la serie de precios de cada una de las criptomonedas. Se ha comprobado si la serie es estacionaria mediante el test de Dickey-Fuller Ampliado. Se decide utilizar este test, ya que este no parte la restrictiva hipótesis de que ε independientes y homocedásticos, sino que este test tiene en consideración que el término de perturbación puede seguir un esquema $u_t \sim AR(p)$. Este test tiene las siguientes hipótesis:

$$\begin{cases} H_0 : Precio_t \sim I(1) \equiv \alpha_1 = 0 \equiv |L| = 1 \equiv \text{proceso no estacionario} \\ H_A : Precio_t \sim I(0) \equiv \alpha_1 < 0 \equiv |L| > 1 \equiv \text{proceso estacionario} \end{cases}$$

Donde L representa la raíz, la cual si es igual a 1 indica que no se puede rechazar que α_1 sea igual a 0, equivalente a decir que no se puede rechazar que sea un proceso integrado de orden 1 que a su vez es equivalente a no poder rechazar que el proceso no es estacionario.

La regresión que se estima es la siguiente:

$$d_{precio_t} = f(t) + \alpha_1 d_{precio_{t-1}} + \sum_{j=1}^{p-1} dd_{precio_{t-j}} + u_t$$

Donde $f(t)$, que es la especificación determinista para el precio de cierre de cada una de las criptomonedas, puede estar expresado de diferentes maneras:

$$\begin{cases} f(t) = Constante \\ f(t) = Constante \text{ ms tendencia} \\ f(t) = Ni constante ni tendencia \end{cases}$$

Podría darse la situación de que el proceso sea integrado de orden 2 y por lo tanto necesite dos diferenciaciones, en este caso las hipótesis y regresión a estimar son las siguientes:

$$\begin{cases} H_0 : d_{Precio_t} \sim I(1) \equiv \alpha_1 = 0 \equiv |L| = 1 \equiv \text{proceso no estacionario} \\ H_A : d_{Precio_t} \sim I(0) \equiv \alpha_1 < 0 \equiv |L| > 1 \equiv \text{proceso estacionario} \end{cases}$$

Que es equivalente a:

$$\begin{cases} H_0 : Precio_t \sim I(2) \equiv \alpha_1 = 0 \equiv |L| = 1 \equiv \text{proceso no estacionario} \\ H_A : Precio_t \sim I(1) \equiv \alpha_1 < 0 \equiv |L| > 1 \equiv \text{proceso estacionario} \end{cases}$$

La regresión que se estima en esta segunda situación es la siguiente:

$$d_{precio_t} = g(t) + \alpha_1 d_{precio_{t-1}} + \sum_{j=1}^{p-1} dd_{Precio_{t-j}} + u_t$$

Donde $g(t)$, que es la especificación determinista para la derivada del precio de cierre de cada una de las criptomonedas, puede estar expresado de diferentes maneras:

$$\begin{cases} g(t) = Constante \\ g(t) = Ni\ constante\ ni\ tendencia \end{cases}$$

Siendo por lo tanto $g(t)$ lo siguiente:

$$g(t) = f(t)' = \Delta f(t)$$

Hasta que este test no ha dado significativo y por lo tanto se ha rechazado la hipótesis nula, no se ha parado de hacer diferenciaciones. Antes de ello, hay diferentes decisiones que hay que tomar a la hora de realizar este test.

17.1.3.1 Determinación del orden de integración de las criptomonedas

Tal y como se ha comentado en el apartado anterior, hace falta determinar que tendencia determinista $f(t)$ tiene cada una de las criptomonedas. Esto se hará haciendo un gráfico de la serie temporal de cada una de las criptomonedas y observando si hay que añadir constante (la serie no se mueve al rededor del 0) o si hay que añadir tendencia (la serie tiene una tendencia ascendente o descendente).

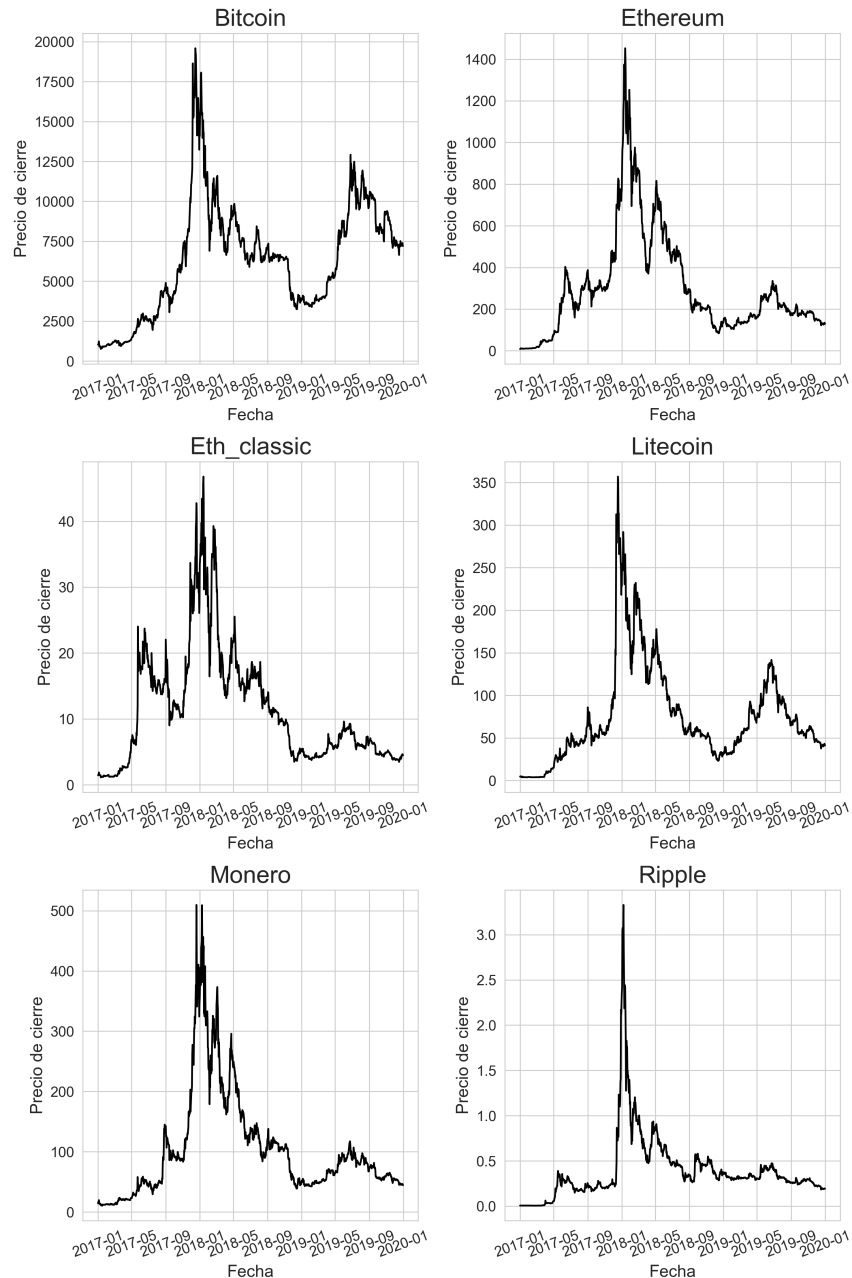


Fig. 27: Serie temporal del precio de cierre para las seis criptomonedas del 1 de enero de 2017 al 31 de diciembre de 2019

Lo primero que se observa en la figura 27 hace falta añadir constante ya el precio no se mueve cerca y al rededor del 0 en ninguna de ellas. En el caso de Ripple el precio se encuentra en niveles que suelen estar en el rango de 0.25 a 1 pero no es 0. Por otro lado se ve como, debido a que todas las criptomonedas están bastante correlacionadas entre ellas, tienen una tendencia ascendente muy importante que empieza al rededor de septiembre de 2017. A partir de ese momento en todas las criptomonedas a excepción de Bitcoin ha habido una tendencia bajista. Aunque hay dos tendencias, una alcista y otra bajista, se decide añadir tendencia a las regresiones de todas las

criptomonedas, ya que se considera que es mejor que no añadir ninguna tendencia. Por lo tanto, todas las criptomonedas tendrán la siguiente especificación:

$$\left\{ \begin{array}{l} f(t) = \text{Constante ms tendencia} \end{array} \right.$$

Determinación del número de retardos y criterio

Antes de calcular el orden de integración de cada proceso hace falta determinar el número de retardos que se añaden, o dicho de otra forma, determinar el parámetro p escogido. Esta elección se hará en base a la estrategia de t-ratio definida en (Ng and Perron 1995). Esta se basa en ir añadiendo retardos hasta que el último retardo añadido tenga un nivel de significación superior al 5%.

17.1.3.2 Aplicación de la diferenciación

Una vez se ha definido como se llevará el cabo el test ADF se analiza el número de veces que hay que diferenciar el precio de cierre de cada criptomoneda para que se rechace la hipótesis nula a un nivel de significación del 5%. El p-valor que determina la significación de α_t siguiendo la aproximación de MacKinnon's explicada en (MacKinnon 1991). MacKinnon crea unas tablas mediante simulación para calcular los valores críticos asintóticos para diferentes tamaños de muestra y diferentes especificaciones deterministas.

Criptomoneda	Estadístico Test	V.C.1%	V.C.5%	V.C.10%	Diferenciaciones
Bitcoin	-6.3080	-3.9672	-3.4146	-3.1295	1
Ethereum	-5.8949	-3.9673	-3.4146	-3.1295	1
Eth_classic	-7.8469	-3.9673	-3.4146	-3.1295	1
Litecoin	-9.2627	-3.9672	-3.4146	-3.1295	1
Monero	-5.9998	-3.9673	-3.4146	-3.1295	1
Ripple	-3.5292	-3.9673	-3.4146	-3.1295	0

Tabla 14: Test ADF para las 6 criptomonedas y número de diferenciaciones aplicadas a cada una de ellas

La tabla 14 muestra como todas las criptomonedas a excepción de Ripple han tenido de ser diferenciadas una vez para conseguir que cada una de las series temporales sea estacionaria. Todas a excepción de Ripple son significativas al 5%. Podría considerarse diferenciar la serie temporal de Ripple, pero como se ha definido un nivel de significación del 5% se mantiene así.

18 Ajuste del modelo ARIMA

En este apartado se ajustan 152 modelos ARIMA a cada una de las criptomonedas atendiendo a la diferenciación d que se ha decidido en el apartado anterior para cada criptomoneda. Para realizar el ajuste de los datos se ha utilizado la librería **statsmodels** de Python. El criterio para escoger el mejor modelo será el criterio de AIC. El modelo ARIMA se puede expresar de esta manera:

$$\phi_p(L)\Delta^d x_t = \theta_q(L)\epsilon_t$$

Donde ϕ_p es el polinomio autoregresivo y $\theta_q(L)$ es el polinomio de medias móviles donde $\epsilon \sim iid(0, \sigma_\epsilon^2)$.

El parámetro Δ^d ya ha quedado definido para cada una de las criptomonedas, teniendo todas $d=1$ a excepción de Ripple, que tiene $d=0$. Faltan por determinar p y q . Para encontrar los mejores p y q , se ha optado por no hacerlo visualmente a través de los correlogramas debido a que se considera que los resultados pueden verse influenciados por la subjetividad de quien los observa, ya que dependiendo del nivel de significación escogido los resultados son unos u otros y no hay ninguna garantía que la elección de la más adecuada para llevar a cabo las predicciones. Es por esto que la forma de proceder ha sido otra.

Como se ha comentado antes, se ha dividido el training en reduced training y validation. Se evalúan todas las combinaciones de q y p para $q \in \{0,1,\dots,n_q\}$ y $p \in \{0,1,\dots,n_p\}$, siendo n_q y n_p dos valores que pueden variar. Alguien inexperimentado podría decidir escoger unas n_p y n_q muy elevadas para así poder abarcar muchos más modelos y escoger el modelo que da el menor AIC, pero esto presenta un problema, y es el sobreajustamiento. Tal y como se ha demostrado en el apartado 7, un modelo subespecificado y que no es una correcta operacionalización de la verdadera función F , puede llegar a tener un error predictivo esperado menor que el modelo del que proceden los datos. Los datos tienen mucho ruido y además no se puede garantizar que el valor absoluto de los parámetros estimados no sea pequeño, por lo que pudiéndose cumplir dos de las condiciones señaladas en (Wu, Harris, and McAuley 2007 citado en Shmueli and others (2010)), esta situación es bastante factible.

Pese a que se usa el criterio de AIC, y tal y como se ha comentado antes, trata de encontrar un equilibrio entre bondad del ajuste y complejidad del modelo, se ha observado que cuando se escogen n_q y n_p elevadas y próximas a 10 hay una tendencia en todas las criptomonedas a que el mejor modelo bajo el criterio de AIC sea aquel más complejo, y por lo tanto con p y q próximas a 10. Este modelo puede ser que esté más cerca de la correcta operacionalización de F , la verdadera función, que otros modelos más simples, pero esto no garantiza mejores predicciones. Es por esto, y dado que el interés que existe es en minimizar el error cuadrático medio en datos no vistos por los modelos, que se decide estudiar el efecto que puede estar teniendo la n_p y n_q máxima en la capacidad predictiva.

Se opta por ajustar todas las criptomonedas en el reduced training set con todas las combinaciones de p y q pero con diferentes rangos de valores de p y q : con n_q y n_p igual a 3, n_q y n_p igual a 5 y n_q y n_p igual a 9 y a continuación se evalúa el error cuadrático medio (ECM) de las predicciones en el validation set. Aquel rango en el que, para cada criptomoneda, se ha obtenido el menor error cuadrático medio, es el escogido para ajustar los datos en el training set y posteriormente evaluar los modelos en el test set. De esta forma se soluciona el problema de que el modelo no sea capaz de generalizar y que se llegue a la conclusión de que las redes LSTM son capaces de predecir mejor los datos futuros cuando el problema real está en el hecho de que en la etapa de ajuste se ha cometido un error que invalida toda conclusión que uno pueda hacer. Para que quede claro, si $n_p=3$ y $n_q=3$, entonces se evalúan para cada criptomoneda 16 modelos, que son los que se generan de hacer todas las combinaciones posibles de p y q .

El mejor enfoque hubiera sido hacer cross validation, dividir el training set en varias partes e ir evaluando la capacidad de predicción en una de las partes, ajustando con el resto y finalmente calcular los promedios para cada uno de los tres rangos. Por ejemplo, si se divide en cinco partes para cada rango se haría un promedio de cinco ECM. Debido al alto coste computacional se opta por el método descrito en el anterior párrafo.

Para hacer las predicciones se usa el método extended de la librería statsmodels de Python. Este consiste en que a la hora de hacer las predicciones en el validation set, las cuales se hacen de una en una y a partir de los precios reales anteriores, no se estiman de nuevo los parámetros con los nuevos datos a medida que se va avanzando en el validation set, sino que por ejemplo, si el modelo ARIMA toma los 10 precios anteriores para hacer la predicción y se está ha prediciendo el elemento 11 del validation set, ninguno de los 10 anteriores se ha incluido al reduced training set para volver a recalcular los parámetros. De esta forma lo que se consigue es no sobreestimar la capacidad predictiva, ya que los parámetros del modelo se estiman solo a partir del reduced training set. Tal y como se verá más adelante, esta metodología de predicción será la misma que se aplicará cuando se usen las redes LSTM.

Rango	X	Bitcoin	Ethereum	Eth_classic	Litecoin	Monero	Ripple
np=3,nq=3	p	2	2	3	3	3	3
	d	1	1	1	1	1	0
	q	3	3	3	2	3	3
	AIC	9810	6371	2502	4685	5364	-1615
	Raíz_ECM	184.101375	7.707319	0.298082	3.125576	3.372207	0.017766
np=5,nq=5	p	4	4	5	4	5	3
	d	1	1	1	1	1	0
	q	4	5	5	4	5	5
	AIC	9802	6362	2484	4670	5323	-1638
	Raíz_ECM	183.971684	7.777363	0.306811	3.201119	3.435348	0.018211
np=9,nq=9	p	9	7	6	6	4	5
	d	1	1	1	1	1	0
	q	7	9	9	7	6	9
	AIC	9768	6337	2484	4657	5323	-1685
	Raíz_ECM	197.924528	8.028963	0.308786	3.302785	3.431120	0.017684

Tabla 15: Raíz del error cuadrático medio (calculado en el validation set) y AIC (calculado en el reduced training) de los mejores modelos de cada rango según el criterio de AIC. Para cada rango y criptomoneda se han encontrado aquellas p y q que minimizan el criterio de AIC. Una vez encontrado el mejor modelo según AIC se evalúa su ECM en el validation set para ver si de los tres modelos con menor AIC encontrados para cada criptomoneda, aquel con menor AIC es el que tiene menor ECM.

Como se puede ver en la tabla 15, donde en total se han evaluado 912 modelos, 152 para cada criptomoneda, la hipótesis de que pese a que el criterio de AIC penaliza la complejidad se estaba produciendo un sobreajustamiento de los datos parece ser cierta. En todas las criptomonedas a excepción de Ripple, la raíz del error cuadrático medio es mayor cuando np y nq son 9 y bajo el criterio de AIC estos eran los mejores modelos, ya que es en estos donde se obtiene el mínimo

valor de AIC para cada criptomoneda. Para Bitcoin, se obtiene una mejor predicción con $np=5$ y $nq=5$ y para el resto de monedas menos Ripple con np y $nq=3$. Se establece el criterio de ajustar cada criptomoneda con el valor de p y q que minimizan la raíz del ECM en el validation set. Este error es más grande dependiendo de la criptomoneda analizada. Esto es debido a que tal y como se ve en la figura 27, cada criptomoneda tiene un rango de precios muy diferente.

Una crítica que podría surgir es que no se ha obtenido el parámetro p y q a través de la observación de las autocorrelaciones simples y parciales. El objetivo del trabajo no es encontrar el modelo ARIMA “real”, el cual tenga todos sus parámetros significativos por motivos como los expuestos en (Shmueli and others 2010), sino aquel que sea capaz de hacer las mejores predicciones; es por ello que lo se valora es el error cuadrático medio en el validation set y no el R^2 o la significación del modelo. Se decide dejar de lado criterios de significación, en los cuales dependiendo del nivel de significación que se defina, del 10%, 5% o 1%, pueden surgir unos modelos u otros con capacidades predictivas muy distintas y se opta por una metodología que permite obtener más garantías de tener un modelo que sea capaz de predecir con bastante precisión.

Una segunda crítica que podría surgir es que al evaluar la capacidad predictiva en el validation set se puede acabar escogiendo el modelo que mejor funciona con esos datos concretos pero no por ello el que mejor generaliza. Al solamente evaluar el ECM en el validation con el mejor modelo para cada rango, o sea tres modelos para cada criptomoneda, esta situación podría darse pero es mucho más improbable que se dé comparado con la situación de haber evaluado el ECM en cientos de modelos. Además, con casi todas las criptomonedas se llega a la conclusión de que con $np=9$ y $nq=9$ las predicciones son peores, por lo que no es algo puntual de una criptomoneda o dos, hecho que hubiera dado apoyo a esa hipotética crítica. Es por ello que se ha optado por tener como primer criterio el de AIC, y a partir de allí escoger el mejor según el criterio de ECM. Otra opción hubiera sido estimar el criterio de AIC consistente, que penaliza más la complejidad que el criterio de AIC, con el que quizá se hubieran obtenido mejores resultados a la hora de generalizar, ya que hubiese habido menos sobreajustamiento.

19 Las Redes neuronales recurrentes vanilla y LSTM

19.1 RNR vanilla

A continuación se presentarán las redes neuronales recurrentes vanilla (RNN) y las redes LSTM (las que se utilizarán en este trabajo). En el camino hasta llegar a las redes LSTM se explicarán muchos conceptos de Machine Learning, los cuales, pese a que requerirían capítulos enteros de muchas páginas cada uno de ellos, se tratarán de explicar lo mejor posible de forma concentrada, para que cualquier lector pueda entenderlo. Además de esto hay una justificación matemática de por qué se acaba usando las redes LSTM en vez de las RNN, la cual, pese a que es complicada de entender, está explicada primero mediante un ejemplo concreto y después de forma generalizada para así facilitar el entendimiento al máximo. Por lo tanto, se ha preferido no presentar la red LSTM y la RNR como una caja negra sino intentar transmitir los potenciales y debilidades que éstas tienen a través del entendimiento de su funcionamiento.

Las redes neuronales recurrentes (RNN) vanilla son redes neuronales que funcionan muy bien para modelar datos secuenciales. La RNN más simple está compuesta por una sola neurona que

recibe una entrada y produce una salida. Esta RNN tiene un mecanismo de bucle que le permite trabajar con datos secuenciales. Primero recibe un input X_0 y tiene como salida un vector h_t o equivalentemente y_t , llamado estado oculto, que permite que la información fluya entre los *time steps*. En caso de unos datos temporales, cada time step t sería un momento del tiempo. En caso de frases, sería cada una de las palabras. Por ejemplo, si quisiéramos evaluar una frase para saber lo que una persona quiere, y esta frase fuera: “qué tiempo hace”, las entradas de cada time step (X_1, X_2, X_3) serían igual a (“qué”, “tiempo”, “hace”), donde cada palabra sería representada en forma de vector de números y y_3 sería la predicción del tiempo que hace.

En la figura 28 se muestra la RNN más simple, compuesta de una sola neurona, que es equivalente a la representación de la figura 29. En ésta se puede observar el mecanismo de bucle que tienen las RNN. En cada time step se produce una salida h_t que se puede utilizar como predicción (flecha vertical hacia arriba) y que sirve como input para el siguiente timestep. Este estado oculto h_t está en función de h_{t-1} y x_t , y se encarga de mantener la información de los inputs anteriores para así, cuando se haga la predicción final (Y_3), dada una secuencia de por ejemplo tres palabras (X_1, X_2, X_3), se disponga información de las dos primeras y por lo tanto que la información pasada no se pierda. En conclusión, el objetivo que se busca con este mecanismo de bucle es que la información de hace n time steps sea tenida en cuenta en la predicción final, pero como se verá más adelante, esto no siempre es así.

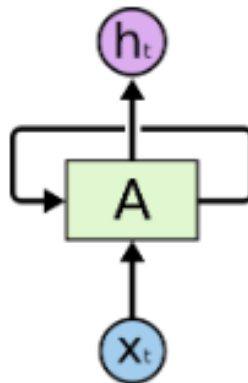


Fig. 28: Neurona de RNN sin desarrollar. Fuente: Understanding RNN and LSTM, Aditi Mittal

En la figura 29 se muestra una neurona como la de la figura 28, pero desarrollada en el tiempo (otra manera de representarla visualmente, por lo tanto, es equivalente a la representación de la figura 28). En el time step 3 la neurona recibe el input X_3 y el estado oculto h_2 del timestep anterior x_{t-1} . En el time step 2 la neurona recibe el input X_2 y da como output h_2 que sirve como input para el time step 3. Esto sucede en todos los time steps, incluido en el primero, donde se suele suponer que recibe una matriz de ceros. Aplicado a este trabajo, este sería un caso donde se trata de predecir el precio de cierre de una criptomoneda (y) en un día a partir de los precios de cierre de los tres días anteriores (X_1, X_2, X_3). Para visualizarlo, sería una fila de la tabla 12 excluyendo la última columna, la cual contiene los precios de cierre que se quieren predecir.

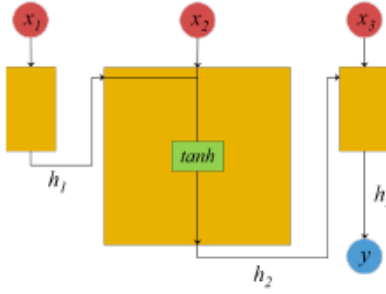


Fig. 29: Neurona desenrollada a través del tiempo con función de activación tangente hiperbólica.
Fuente: (Ji, Kim, and Im 2019)

Dado un *input* X_t , cada estado oculto se calcula de esta forma (Ji, Kim, and Im 2019) :

$$h_t = \tanh(W_h x_t + U_h h_{t-1} + b_h) \quad (3)$$

donde W_h , U_h y b_h son los parámetros que la RNN estima. **Estos tres parámetros son los mismos en todos los time steps.** A diferencia de con el uso de redes neuronales profundas, para calcular y_t no solo se usa el input X_t sino que se usa el estado oculto anterior h_{t-1} (tal y como se observa en la fórmula 3), que a su vez depende de los estados anteriores. De esta forma la red neuronal recurrente consigue explotar la dependencia temporal de las secuencias como se han comentado anteriormente.

La predicción del precio de cierre y_t basada en el precio de los tres días anteriores queda, por tanto, definida de esta manera (Ji, Kim, and Im 2019):

$$y = w_y h_3 + b_y \quad (4)$$

Donde w_y y b_y también son parámetros que la RNN tendrá que aprender. Este proceso de ir calculando los estados ocultos hasta llegar a la/las predicciones finales se define como propagación hacia adelante. Como se puede ver en la fórmula 4, la predicción del precio y se hace a partir del estado oculto h_3 . Este estado oculto a su vez depende de los dos anteriores (fórmula 3), y por lo tanto, la predicción tiene en consideración todos los inputs.

Tanh es la función tangente hiperbólica, la cual permite que el output sea acotado al rango de $[-1,1]$ y permite que el output final (la predicción) dependa de forma no lineal del input y que por lo tanto que se puedan patrones muy complejos de los datos. Además, su segunda derivada puede mantenerse durante bastante tiempo sin ir a cero, propiedad importante como se verá más adelante. Ésta función viene expresada de la siguiente manera:

$$\tanh x = \frac{1 - \exp(-2x)}{1 + \exp(-2x)}$$

Donde x es lo que hay entre paréntesis en la fórmula 3.

Su forma gráfica y la de otras funciones de activación que se explicarán más adelante, junto a sus derivadas, se pueden ver en la figura 30. Esta función de activación y el resto, a excepción de la *step*, tienen la principal función de introducir no linealidad y de esta forma hacer que el output dependa no linealmente del input, pudiendo así captar patrones de los datos más complejos. En la figura se puede ver que la función tangente hiperbólica tiene un rango que va de -1 a 1, esto significa que el estado oculto de una RNN siempre se encuentra en este rango si la función de activación es la tanh.

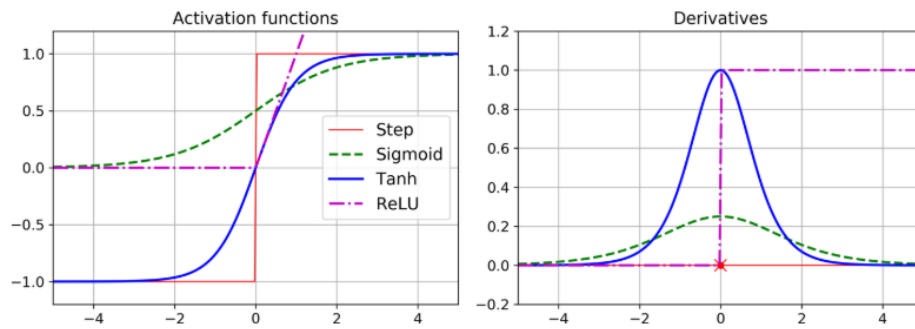


Fig. 30: Función de activación tangente hiperbólica. Fuente: (Géron 2019)

Lo que la RNN definida en este trabajo tratará de minimizar para cada secuencia de días es el error cuadrático medio (ECM), que se define como $E(Y_1, Y_2, \dots, Y_T)$, donde las Y_t son las predicciones en el time step t y T es el time step máximo. En este trabajo, como solo se hace una sola predicción para cada secuencia y ésta se hace en el último time step, aunque la RNN te da el resultado de todas las predicciones, solo Y_T se tendrá en cuenta para calcular el error mediante el ECM y posteriormente propagarlo en la llamada fase de retropropagación hacia atrás, que será explicada a continuación. Entonces, según cual sea la m definida, que en este caso será igual a T , y mediante el ECM, se calculará el error en la predicción del elemento $m+1$ de cada fila de la tabla S normalizada. Como solo se hace la predicción de un día, calcular el ECM es equivalente a calcular el error cuadrático. Para que se entienda la manera de expresar el coste, si la m escogida es 3 y por lo tanto, las predicciones se hacen a partir de los precios de cierre de los tres días anteriores, entonces el coste se define como E_3 .

19.1.1 Método del gradiente descendiente

El método más utilizado para actualizar los parámetros es el del gradiente descendiente. Éste parte de la base de que la función que se quiere minimizar es diferenciable. El método que se usará en el trabajo es el de Adam, el cual es bastante más complejo, pero hace falta entender este para comprender el funcionamiento del de Adam.

Tomando la función $f(x)$, siendo x un valor concreto en el que se evalúa la función, el gradiente de la función es la tasa de cambio de $f(x)$ cuando se cambia el parámetro x , la cual puede aproximarse como $\frac{\partial f}{\partial x}$. Si lo que se busca es cambiar el valor de $f(x)$ se deberá cambiar el valor de x , encontrándonos en 3 casos:

- Que $\frac{\partial f}{\partial x} < 0$ y por lo tanto que $f(x)$ disminuye a medida que aumenta x .

- Que $\frac{\partial f}{\partial x} > 0$ y por lo tanto que $f(x)$ aumenta a medida que aumenta x .
- Que $\frac{\partial f}{\partial x} = 0$ y por lo tanto $f(x)$ está en un mínimo o máximo ya sea local o no.

El método del gradiente descendiente actualiza el valor de x hasta llegar a la situación en la que $\frac{\partial f}{\partial x} = 0$. El método para actualizar el valor de x es el siguiente:

$$x_{nueva} = x_{antigua} - \alpha \frac{\partial f(x)}{\partial x} \quad (5)$$

Donde α es la tasa de aprendizaje, la cual toma valores entre 0 y 1, y se puede optar por mantenerla fija o variarla a lo largo del aprendizaje. En la figura 31 se ve un ejemplo de aplicación con la función $f(x)=x^2$. Si se empieza en $x=-2$, entonces $\frac{\partial f}{\partial x}$ es <0 y por lo tanto como en la expresión 5 el gradiente se está restando, entonces el valor de x_{nueva} actualizado aumentará hasta situarse cerca de 0. En caso de que se hubiera empezado el procedimiento del gradiente descendiente con $x=2$, el valor de $\frac{\partial f}{\partial x}$ hubiera sido >0 , por lo que al restarlo en la expresión (1) el valor de x_{nueva} hubiera disminuido con lo que también hubiera habido un acercamiento progresivo al óptimo que es $x=0$.

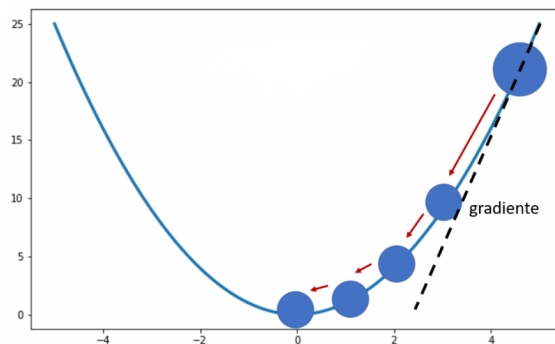


Fig. 31: Gradiente descendiente aplicado a la función $y=x^2$ con una tasa de aprendizaje baja. El eje horizontal representa los valores que toma x , el eje vertical los valores de $f(x)$. Fuente: Gradiente Descendiente para aprendizaje automático, Jose Martinez Heras

Hay una aspecto a tener muy en cuenta, y es que si se escoge una tasa de aprendizaje α muy elevada se corre el riesgo de que pase algo como lo de la figura 32 y que por lo tanto no se llegue al óptimo. Es por ello que se suelen escoger tasas de aprendizajes próximas a 0.

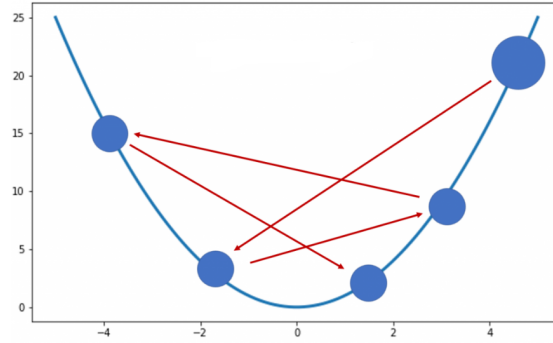


Fig. 32: Gradiente descendiente aplicado a la función $y=x^2$ con una tasa de aprendizaje elevada. El eje horizontal representa los valores que toma x , el eje vertical los valores de $f(x)$. Fuente: Gradiente Descendiente para aprendizaje automático, Jose Martinez Heras

19.1.2 Formulación de la fase de retropropagación a través del tiempo para caso concreto

La retropropagación es la fase en la cual se lleva a cabo la actualización de los parámetros. El ejemplo de la figura 29 consiste en una sola neurona desarrollada a través del tiempo, en el que se predice un valor que en este caso es Y_3 a partir de los precios de cierre de los tres días anteriores (X_1 , X_2 y X_3). Con Y_3 es con el que se calcula el coste E_3 mediante la fórmula del ECM.

Dicho informalmente, la red recibe como input los valores de un individuo, mediante los parámetros existentes hasta ese momento del entrenamiento (W_h , U_h , b_h , etc.). con esos parámetros lleva a cabo una predicción y es a partir de esta predicción que se calcula el coste. Con este coste se calcula el gradiente y por lo tanto se hacen variar los parámetros en el sentido que permita que la función de coste se reduzca. La red realiza la predicción del siguiente individuo no con el valor de los parámetros anteriores, sino con los nuevos parámetros actualizados por el anterior individuo. Este proceso sigue hasta que se empieza a correr el riesgo de sobreajustar. Este sobreajustamiento se puede ver gráficamente evaluando el error en el validation set en cada momento del entrenamiento. Si éste se incrementa en vez de reducirse a la vez que el error en el training set, con el que se está ajustando el modelo, no para de reducirse, entonces se está produciendo un sobreajustamiento, ya que se está perdiendo la capacidad de generalización del modelo (cada vez predice peor en datos que no ha visto).

En la figura 33 se muestra la fase propagación hacia adelante explicado en los párrafos anteriores (flechas negras) y la fase de retropropagación representada por las flechas rojas. En esta figura las "s" hacen referencia a los estados ocultos h_t , y por lo tanto es una forma equivalente de escribirlos. El objetivo de la fase de retropropagación es ir actualizando los parámetros W_h , U_h y b_h de la fórmula 3 y los parámetros W_y y b_y de la fórmula 4. Se va calculando la derivada del coste E_3 respecto a estos cinco parámetros y se hace que estos varíen en el negativo del gradiente, consiguiendo así cambiarlos para que minimicen el coste y así conseguir mejores predicciones. Es E_3 ya que en este ejemplo se calcula el coste a partir de los tres precios anteriores y por lo tanto $T=3$. En este apartado se profundizará en los casos de W_h y U_H , ya que son los que permiten concluir el por qué es mejor evitar el uso de las RNN.

Las fórmulas que permiten el cálculo de los tres estados ocultos son las siguientes. Tal y como se

ha dicho, W_h , U_h , y b_h son iguales en todos los time steps y por lo tanto en el cálculo de los tres estados ocultos tienen el mismo valor.

$$S_1 = h_1 = \tanh(W_h x_1 + U_h h_0 + b_h)$$

$$S_2 = h_2 = \tanh(W_h x_2 + U_h h_1 + b_h)$$

$$S_3 = h_3 = \tanh(W_h x_3 + U_h h_2 + b_h)$$

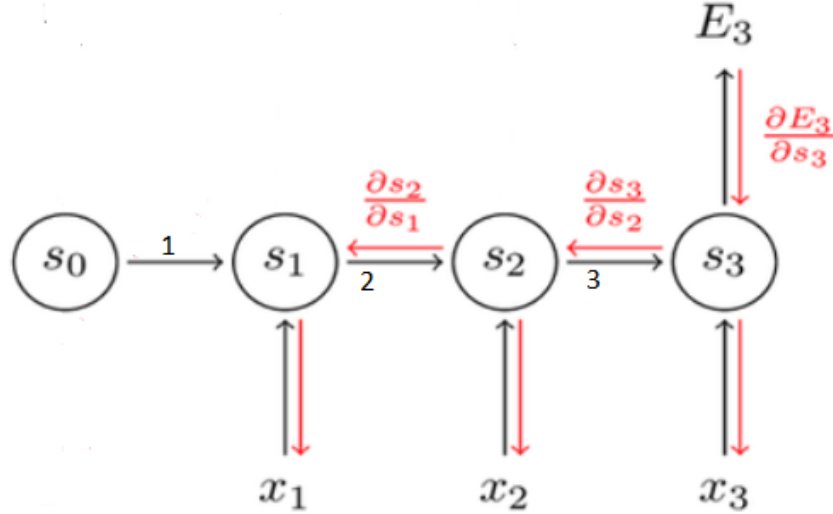


Fig. 33: Ejemplo de retropropagación con $m=3$. Las flechas negras y representan los cálculos que se hacen en la fase de propagación hacia adelante, las flechas rojas los cálculos que se hacen en la retropropagación para actualizar los parámetros

Primero de todo, tal y como se puede observar en la fórmula 6 y 7 de abajo, para calcular el gradiente del error respecto a W_h y U_h se ha calculado la derivada de E_3 respecto a estos parámetros tantas veces como time steps hay, es decir, tres veces. Simplificando la suma de los tres elementos que hay en estas dos fórmulas se podría llegar a pensar que se está calculando el mismo número tres veces, pero no es así. El primer sumando de los tres hace referencia a la derivada del coste E_3 respecto al parámetro W_h que se utiliza en el cálculo de S_1 o equivalentemente h_1 señalado con un 1 en la figura 33. Este parámetro afecta por lo tanto a h_2 y h_3 (tal y como se ha visto en la fórmula 3). Es por por ello que el efecto de variar W_h en el primer sumando es distinto al del segundo sumando, que representa lo mismo pero para el parámetro W_h que se usa en el cálculo de h_2 . Como se observa, cuanto más lejos se está del cálculo de E_3 más multiplicaciones se producen aplicando la regla de la cadena y por lo tanto esto implica que la influencia que tienen los time steps más lejanos en la actualización de los parámetros (primer elemento del sumando) es mucho menor a la que tienen los más cercanos (tercer elemento del sumando). Por ahora es algo que se puede intuir y que se mostrará al final de este apartado.

Cabe destacar que S_0 o equivalentemente h_0 se suele inicializar en 0 tal y como se menciona en (Géron 2019).

$$\frac{\partial E_3}{\partial W_h} = \sum_{t=1}^3 \frac{\partial E_3}{\partial y_3} \frac{\partial y_3}{\partial h_3} \frac{\partial h_3}{\partial h_t} \frac{\partial h_t}{\partial W_h} = \frac{\partial E_3}{\partial y_3} \frac{\partial y_3}{\partial h_3} \frac{\partial h_3}{\partial h_2} \frac{\partial h_2}{\partial h_1} \frac{\partial h_1}{\partial W_h} + \frac{\partial E_3}{\partial y_3} \frac{\partial y_3}{\partial h_3} \frac{\partial h_3}{\partial h_2} \frac{\partial h_2}{\partial W_h} + \frac{\partial E_3}{\partial y_3} \frac{\partial y_3}{\partial h_3} \frac{\partial h_3}{\partial W_h} \quad (6)$$

El mismo procedimiento se sigue con U_h .

$$\frac{\partial E_3}{\partial U_h} = \sum_{t=1}^3 \frac{\partial E_3}{\partial y_3} \frac{\partial y_3}{\partial h_3} \frac{\partial h_3}{\partial h_t} \frac{\partial h_t}{\partial U_h} = \frac{\partial E_3}{\partial y_3} \frac{\partial y_3}{\partial h_3} \frac{\partial h_3}{\partial h_2} \frac{\partial h_2}{\partial h_1} \frac{\partial h_1}{\partial U_h} + \frac{\partial E_3}{\partial y_3} \frac{\partial y_3}{\partial h_3} \frac{\partial h_3}{\partial h_2} \frac{\partial h_2}{\partial U_h} + \frac{\partial E_3}{\partial y_3} \frac{\partial y_3}{\partial h_3} \frac{\partial h_3}{\partial U_h} \quad (7)$$

Una vez calculado esto se pueden actualizar los parámetros mediante el método del gradiente descendiente. Este consiste en restar el gradiente del error ($\frac{\partial E_3}{\partial W_h}$ al parámetro W_h y $\frac{\partial E_3}{\partial U_h}$ al parámetro U_h) para así hacerlos variar en la dirección que minimiza el coste. Se usa un parámetro α el cual puede ir cambiando en el entrenamiento o se puede mantener fijo, que sirve para definir en cuanta magnitud se desea que varíen los parámetros cuando se aplica retropropagación.

$$W_h = W_h - \alpha \frac{\partial E_3}{\partial W_h} \quad (8)$$

$$U_h = U_h - \alpha \frac{\partial E_3}{\partial U_h} \quad (9)$$

Se puede entender el gradiente como el indicador de hacia que sentido hay que mover cada parámetro para minimizar una función de cost determinada (si resta el gradiente, en caso de no cambiarle el signo indica el sentido para maximizar la función). Cada individuo, al cometer un error distinto en la predicción, dirá un sentido distinto en el que hay que mover los parámetros, pero si se va iterando sobre el training set se acaban teniendo unos parámetros que permiten estar cerca del óptimo. Cuanto de cerca se quede del óptimo depende de muchas variables, entre ellas el método utilizado para actualizar los parámetros de la fórmula (8) y (9), ya que existen los siguientes:

- *stochastic gradient descent*, que consiste en ir actualizando los parámetros del modelo utilizando las filas de la tabla 12 (a excepción de la última columna) de una en una, por lo que el error que se propaga en la fase de retropropagación en el tiempo es el error cometido en la predicción del precio de un día.
- otra alternativa es *mini-batch gradient descent* que consiste en utilizar grupos de n individuos para actualizar los parámetros, calculando el promedio de los n gradientes y actualizando los parámetros del modelo con ese promedio (es como hacer *stochastic gradient descent* con n individuos, solo que en vez de actualizar los parámetros n veces, se actualizan una vez a través del promedio de los gradientes). Como se puede intuir, al estar actualizando en base a varios individuos y no solo uno, la convergencia al óptimo, ya sea local o global, es bastante menos errática que la del primero.
- finalmente con el *batch gradient descent* se utilizan todos los individuos (p elementos) del training para, mediante el promedio de los p gradientes, actualizar los parámetros.

En la figura 34 se puede observar un caso en el que se requiere obtener el valor óptimo de dos parámetros θ_0 y θ_1 . En el caso de batch gradient descent, cada círculo es un epoch (actualización de los parámetros mediante el uso de todos los individuos del training), en Mini-batch gradient descent cada símbolo de suma es una actualización de los parámetros mediante los individuos de cada grupo de n elementos y finalmente cada cuadrado de Stochastic gradient descent es una actualización de los parámetros mediante un individuo. Mientras que con batch gradient descent se tiende al óptimo, con las otras dos metodologías se acaba dando vueltas alrededor del mismo, aunque como se puede ver, con mini-batch gradient descent se suele estar más cerca.

En este trabajo se ha optado por utilizar grupos de 20 elementos (batch size = 20) para actualizar los parámetros. Se ha observado que mediante el uso de uno solo (batch size= 1) el error resultante en el validation set era mucho mayor. Esto es resultado de lo que se muestra en la figura 34.

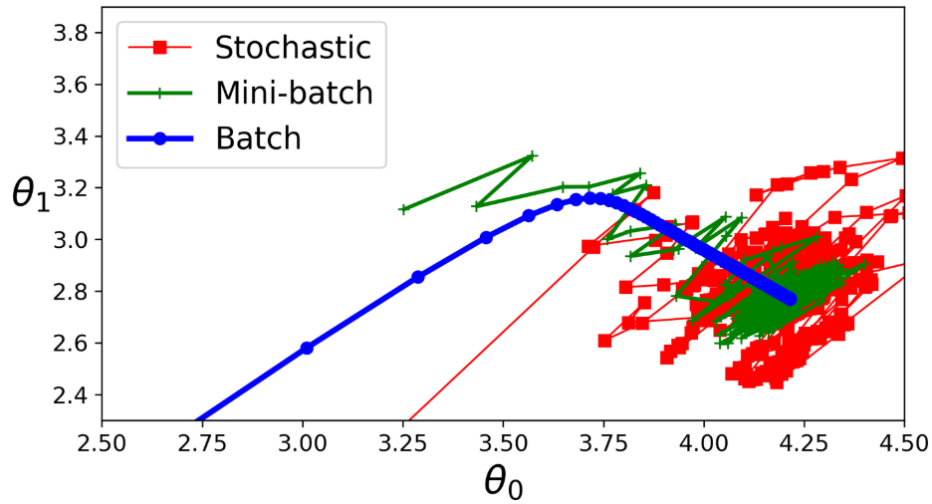


Fig. 34: Gradient Descent aplicado con tres metodologías para la resolución de un problema de regresión con dos parámetros. Fuente: [@geron2019hands]

Para aquellos que quieran entrar en más detalle, se ha utilizado el optimizador de Adam (Kingma and Ba 2014), por lo que el método usado se podría llamar *Mini-Batch adam optimization*. El optimizador Adam en vez reducir el gradiente a un ritmo elevado pero llevando mucho tiempo de entrenamiento para llegar al óptimo, se va acercando de forma acelerada hasta que alcanzarlo. Esto lo hace a través de dos momentos (“m” y “s”), los cuales se van acelerando a medida que avanza el entrenamiento. Estos dos momentos contienen el gradiente actual. La diferencia está en que el momento “m” definido bajo este párrafo contiene un promedio exponencial decadente de los anteriores gradientes. El gradiente “s” es igual pero tiene los gradientes anteriores elevados al cuadrado. No se va a entrar en detalles sobre el mismo, pero quién tenga curiosidad aquí está su formulación de forma genérica (Kingma and Ba 2014):

$$1. m = \beta_1 m - (1 - \beta_1) \Delta_{\theta} J(\theta)$$

$$2. s = \beta_2 s - (1 - \beta_2) \Delta_{\theta} J(\theta) \otimes \Delta_{\theta} J(\theta)$$

$$\begin{aligned}\hat{m} &= \frac{m}{1 - \beta_1^T} \\ \hat{s} &= \frac{s}{1 - \beta_2^T} \\ \theta &= \theta + \alpha \hat{m} \oslash \sqrt{(\hat{s} + \epsilon)}\end{aligned}$$

Donde:

$\Delta_{\theta} J(\theta)$ es el gradiente

Además, $J(\theta)$ es la función de coste, \otimes es la multiplicación por elementos, \oslash es la división por elementos, θ son los parámetros que se están actualizando, y m y s son los momentos.

19.1.3 Formulación de la fase de retropropagación a través del tiempo generalizada

Una vez se ha explicado detalladamente y para un ejemplo concreto el procedimiento de propagación hacia adelante y retropropagación a través del tiempo aplicado al caso de las redes neuronales recurrentes más simples que existen, se procede a explicar, tal y como se hace en (Pascanu, Mikolov, and Bengio 2013), por qué las RNN tienen serios problemas con series temporales largas, debido a dos situaciones que se pueden dar: que el gradiente se desvanezca a medida que la serie temporal se alarga y que por lo tanto que el input más lejano no sea tenido en cuenta o que pase todo lo contrario y a medida que se propaga el error a través de más time steps, el gradiente se va haciendo cada vez más grande, también conocido como *exploding gradient*. Es por esto que se optará por utilizar las redes LSTM que serán explicadas más adelante.

Siguiendo la formulación de la fórmula (6), se desarrolla el cociente siguiente:

$$\frac{\partial h_3}{\partial h_t}$$

La diferencia es que ahora es para cualquier time step máximo T , quedando por lo tanto así:

$$\frac{\partial h_T}{\partial h_t}$$

El caso mostrado en la fórmula (6) y (7) era para el time step 3 ($T=3$), ya que la longitud de la secuencia era 3. Ahora se generaliza con un tamaño de secuencia T .

$$\frac{\partial h_T}{\partial h_t} = \frac{\partial h_T}{\partial h_{T-1}} \frac{\partial h_{T-1}}{\partial h_{T-2}} \dots \frac{\partial h_{t+1}}{\partial h_t} = \prod_{k=t}^{T-1} \frac{\partial h_{k+1}}{\partial h_k} \quad (10)$$

A partir del cociente $\frac{\partial h_{k+1}}{\partial h_k}$ se llega a esta expresión (Pascanu, Mikolov, and Bengio 2013):

$$\frac{\partial h_{k+1}}{\partial h_k} = \text{diag}(\tanh'(W_h x_t + U_{h_{t-1}} + b_h)) U_h \quad (11)$$

donde diag convierte un vector en una matriz diagonal y donde W_x es una matriz $n_{inputs} \times n_{neuronas}$. Si se utiliza el precio de cierre, de apertura y el máximo para predecir el precio de

cierre entonces el número de inputs son 3. Se ha de tener en cuenta que esta derivada parcial recursiva que se está calculando es una matriz jacobiana.

Si se propaga a través de k time steps entonces el gradiente es el siguiente (Pascanu, Mikolov, and Bengio 2013):

$$\frac{\partial h_{k+1}}{\partial h_1} = \prod_t^K \text{diag}(\tanh'(W_h x_t + U_h h_{t-1} + b_h)) U_h \quad (12)$$

Si el valor propio más grande de la matriz U_h es más grande que 1, entonces el gradiente crece descontroladamente y el producto se hace muy grande, por lo contrario, si es menor que 1, el gradiente se desvanece, tal y como se concluye en (Pascanu, Mikolov, and Bengio 2013). Para elaborar este trabajo no se ha entrado en detalle en esta demostración, ya que no se ha considerado necesario. Cabe destacar, y tal y como se muestra expresamente en la figura 30, la derivada de la función de activación tangente y del resto de funciones de activación es siempre menor o igual a 1 y por lo tanto, si el valor propio más elevado es menor a 1 entonces la derivada de la fórmula 11 tenderá a 0. Esta derivada nos informa sobre el efecto que tiene un cambio en el estado oculto k cuando se produce un pequeño cambio en el estado oculto que está en el timestep 1. Esto implica que si el gradiente se desvanece, entonces los estados ocultos más lejanos no tendrán ningún efecto en los posteriores, y en conclusión, no se aprenden las dependencias a largo plazo. Dicho de otra forma y aplicado a este trabajo, lo que pasó hace 10 días si $m=10$ y si existe el problema del gradiente descendiente, no afectará a los estados ocultos posteriores y por lo tanto no afectará prácticamente nada en la actualización de los parámetros, o dicho de otra forma, la actualización de los parámetros no tomará en consideración su estado oculto.

19.2 Red LSTM

19.2.1 Estructura de la celda LSTM

La Long Short Time Memory o LSTM es una arquitectura de red neuronal recurrente artificial utilizada en el campo del aprendizaje profundo. Hay un conjunto de operaciones dentro de las neuronas que permiten a la LSTM mantener u olvidar información. Por lo tanto, las LSTM comparado con las redes recurrentes estándar, tienen la capacidad de aprender dependencias largas ya que tienen una memoria a más largo plazo. Consiste también en una cadena de celdas pero estas tienen una estructura más compleja que las celdas de una red neuronal recurrente. En estas celdas en vez de haber una única capa de red neuronal (la tangente), hay cuatro, las cuales interactúan de una forma especial.

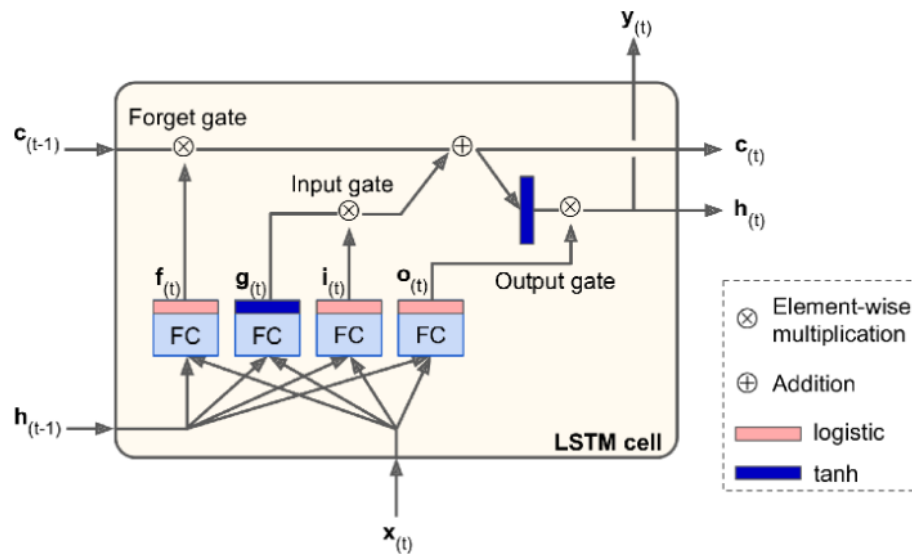


Fig. 35: Celda LSTM. Fuente: (Géron 2019)

En la figura 35 se puede ver la estructura interna de una celda LSTM. Los rectángulos rosas indican que se aplica la función logística a los inputs y los azules que se aplica la función tangente.

La función de activación logística está definida de la siguiente manera:

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

Tiene la particularidad de que si aplicas esta función a cualquier input la salida siempre está en un rango de 0 a 1, a diferencia de la función tangente hiperbólica que tiene un rango que va de -1 a 1.

A diferencia de la celda de una RNN que se ha mostrado en la figura 29, esta posee muchos más elementos. Lo primero de todo que cabe destacar es que en vez de tener un solo estado como en el caso de la RNN, tiene dos estados:

- El estado a largo plazo es C_t , como su nombre indica, mantiene la información a lo largo del tiempo.
- h_t es el estado a corto plazo, que mantiene la información de time steps cercanos.

La clave es que a medida que el estado a largo plazo atraviesa la celda, primero pasa por lo que se llama la compuerta del olvido o *forget gate*, en la cual se elimina cierta información, para a continuación, en la operación aditiva añadir nueva información que ha sido seleccionada por la compuerta de entrada o *input gate*. Entonces c_t se le da un doble uso:

- Sale directamente de la celda sin ser modificado.
- Se copia y es pasado a través de una función tangente hiperbólica y el resultado es filtrado por la compuerta de salida o *output gate*. Esto produce lo que se denomina como estado a corto plazo o h_t que es igual al output de la celda y_t .

Primero de todo el vector de input $X_{(t)}$ y el estado a corto plazo $h_{(t-1)}$ alimentan a cuatro capas. Cada una tiene su función, tal y como se indica en (Géron 2019):

- $g_{(t)}$: en la celda que se ha analizado de la RNN solo existe esta capa, y como se ha visto, tiene el rol de analizar el input $X_{(t)}$ y el estado oculto previo $h_{(t-1)}$ y el output de esta operación sale directamente. En la celda LSTM en vez de eso tiene la función de añadir al estado de largo plazo la información más importante, y el resto se elimina.
- Las otras tres capas, coloreadas en rosa en la figura 35, tienen la función de ser controladores de compuerta. Como utilizan una función sigmoide o logística, el output va de 0 a 1. En la figura 35 se puede observar como el output de estas capas va a parar a operación de multiplicación por elementos, y por lo tanto si la salida es un 0, cierran la compuerta. En el caso de la compuerta de olvido, controlada por $f_{(t)}$, controla que información se olvida. Para la compuerta de entrada, controlada por $i_{(t)}$, controla que partes de $g_{(t)}$ deben añadirse al estado de largo plazo. Finalmente $o_{(t)}$ controla que partes del estado a largo plazo deben tenerse en cuenta y han de formar parte de las salidas $h_{(t)}$ y $y_{(t)}$.

Las computaciones son las siguientes (Géron 2019):

1. $i_{(t)} = \sigma(W_{xi}^T X_{(t)} + W_{hi}^T h_{(t-1)} + b_i)$
2. $f_{(t)} = \sigma(W_{xf}^T X_{(t)} + W_{hf}^T h_{(t-1)} + b_f)$
3. $o_{(t)} = \sigma(W_{xo}^T X_{(t)} + W_{ho}^T h_{(t-1)} + b_o)$
4. $g_{(t)} = \tanh(W_{xg}^T X_{(t)} + W_{hg}^T h_{(t-1)} + b_g)$
5. $C_{(t)} = f_{(t)} \otimes C_{(t-1)} + i_{(t)} \otimes g_{(t)}$
6. $y_{(t)} = h_{(t)} = o_{(t)} \otimes \tanh(c_t)$

En estas ecuaciones:

- $W_{xi}, W_{xf}, W_{xo}, X_{xg}$ hacen referencia a las matrices de parámetros de cada una de las cuatro capas para su conexión con el vector de input x_t .
- $W_{hi}, W_{hf}, W_{ho}, X_{hg}$ hacen referencia a las matrices de parámetros de cada una de las cuatro capas para su conexión con el vector de input h_{t-1} .
- b_i, b_f, b_o y b_g son los términos de sesgo de cada una de las capas.

19.2.2 Por qué las LSTM solucionan el problema del gradiente desvaneciente

Como ya se ha visto, la derivada recursiva provocaba que el gradiente se desvaneciese haciendo que la RNN no pudiese aprender dependencias de largo plazo. Las LSTM solucionan esto con la ayuda del estado a largo plazo $C_{(t)}$. Originalmente, y tal y como se señala en ("Why LSTMs Stop Your Gradients From Vanishing: A View from the Backwards Pass", s.f., párrafos 12-13), la ecuación para actualizarlo era esta:

$$C_{(t)} = C_{(t-1)} + i_{(t)} \otimes g_{(t)}$$

Haciendo que $C_{(t)}$ dependiese de su previo valor y añadiéndole el resultado de la multiplicación por elementos de i_t y $g_{(t)}$. Debido a que el valor de $C_{(t)}$ tendía a aumentar se añadió la controladora de compuerta $g_{(t)}$ para escalar el estado a largo plazo, finalizando con la formulación 13, que es la expuesta en el anterior apartado:

$$C_{(t)} = f_{(t)} \otimes C_{(t-1)} + i_{(t)} \otimes g_{(t)} \quad (13)$$

Para ver por qué ahora no existe el problema del gradiente desvaneciente causado por la derivada recursiva, se expande, mediante la regla de la cadena multivariante, la derivada completa de:

$$\frac{\partial C_t}{\partial C_{t-1}}$$

$$\frac{\partial C_t}{\partial C_{t-1}} = \frac{\partial C_t}{\partial f_t} \frac{\partial f_t}{\partial h_{t-1}} \frac{\partial h_{t-1}}{\partial C_{t-1}} + \frac{\partial C_t}{\partial i_t} \frac{\partial i_t}{\partial h_{t-1}} \frac{\partial h_{t-1}}{\partial C_{t-1}} + \frac{\partial C_t}{\partial g_t} \frac{\partial g_t}{\partial h_{t-1}} \frac{\partial h_{t-1}}{\partial C_{t-1}} + \frac{\partial C_t}{\partial C_{t-1}} \quad (14)$$

Escribiendo explícitamente las derivadas queda así:

$$\begin{aligned} \frac{\partial C_t}{\partial C_{t-1}} = & c_{t-1} \sigma'(\hat{u}) W_f * o_{t-1} \tanh'(C_{t-1}) + g_t \sigma'(\hat{u}) W_i * o_{t-1} \tanh'(C_{t-1}) + \\ & + i_t \tanh'(\hat{u}) W_c * o_{t-1} \tanh'(C_{t-1}) + f_t \end{aligned} \quad (15)$$

La diferencia, tal y como se afirma en (“Why LSTMs Stop Your Gradients From Vanishing: A View from the Backwards Pass”, s.f., párrafos 18-20), está en que si se quiere hacer retropropagación en el tiempo k time steps hacia atrás, sencillamente hay que añadir los términos de la fórmula (14) k veces. La diferencia principal es que mientras que con las RNN mostrada anteriormente, en una situación como esta, el gradiente acababa o muy por encima de 1 o tendía hacia 0. Ahora, con la expresión de la fórmula (15), esta derivada parcial en cada timestep puede tomar cualquier valor en el rango $[0,1]$ o superior a 1, pero no tiene por qué converger a 0. En el supuesto de que esto pasara, tal y como indica en este artículo, siempre se puede dar valores superiores a los controladores de compuerta. Cabe destacar que los valores de f_t , o_t , i_t y g_t serán aprendidos por la red neuronal a lo largo del entrenamiento, y por lo tanto aprenderá cuando preservar o hacer desaparecer el gradiente.

19.2.3 Ajuste de la red LSTM

Con tal de ajustar la red LSTM se ha hecho es apilar varias capas de celdas LSTM, para entender qué es esto, la figura 36 muestra un ejemplo en el que se lleva a cabo la predicción de un precio en base al precio de los tres días anteriores, y en la cual hay dos celdas LSTM extendidas a través del tiempo (horizontalmente).

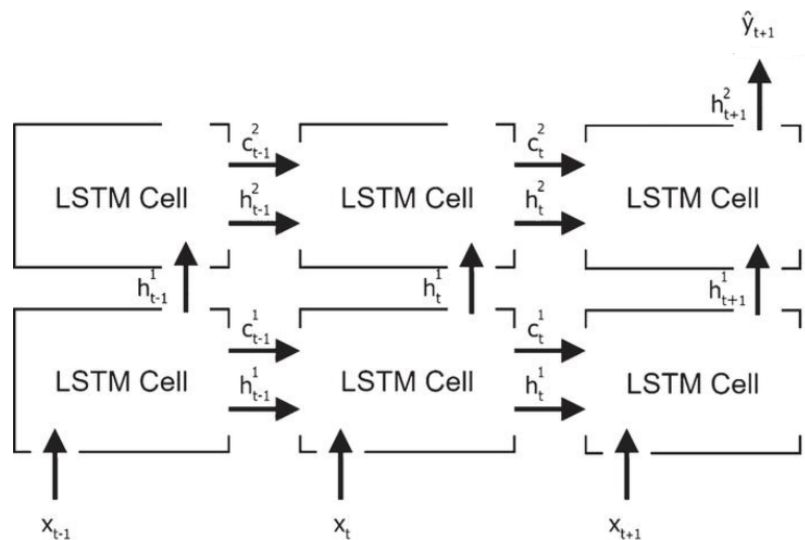


Fig. 36: Dos celdas LSTM apiladas para la predicción de un precio en base al precio de los tres días anteriores. Fuente: (Nicholas et al. 2018)

Como se puede ver en la figura 36, el estado a largo plazo c_{t-1} y el estado a corto plazo h_{t-1} se transmite entre time steps para una misma celda LSTM y lo que se transmite de celda a celda en cada timestep es el estado a corto plazo. Encima de cada estado hay un número indicando la celda a la que pertenece. Así, por ejemplo, el estado h_t^1 que se encuentra en el centro de la figura es aquel que se transfiere de la celda 1 a la celda 2 en el timestep X_t . La predicción final es \hat{y}_{t+1} y es a partir de la cual se lleva el procedimiento de retropropagación, solo que ahora el número de computaciones es superior al visto hasta ahora, ya que ahora no solo hay una neurona.

Para este trabajo se ha escogido que en la capa intermedia haya tantas neuronas o equivalentemente celdas LSTM como el tamaño de m . De tal forma que si se quiere predecir el precio a partir de los 15 días anteriores habrá 15 neuronas aplicadas una encima de la otra. De esta forma lo que se consigue es dos cosas:

- Aumentar la complejidad de la red LSTM a medida que la tarea es más complicada, ya que se requiere de aprender de patrones más complicados cuanto mayor es la m .
- Evitar sobreajustar los datos. Un problema que se ha visto que se tenía y que se ha corregido es que al principio, usando 50 neuronas, el modelo conseguía tener un error cuadrático medio evaluado en el reduced training set muy bajo, pero una vez evaluado en el validation se obtenían mucho peores resultados comparado con la estrategia de usar tantas neuronas como el valor m . Es una práctica generalizada la de elegir un número de neuronas que esté comprendido entre el número de neuronas de input (m) y el número de neuronas de output (1 ya que solo se predice el precio del día siguiente). Es por esto, y con tal de evitar un sobreajustamiento, que se ha optado por esta estrategia. Hay otros que utilizan unas fórmulas para el cómputo del número óptimo de neuronas, pero estas requieren encontrar

un valor adecuado para un parámetro de la misma. Visto el añadido de complejidad que esto hubiera supuesto, se ha optado por la estrategia mencionada.

Antes de ajustar en todo el training set la red LSTM se ha decidido, como en el caso del modelo ARIMA, utilizar un validation set para encontrar la estructura de la red LSTM que mejor se adapta a cada criptomoneda. Se ha estudiado como el tamaño de m , o sea, el número de días en base a los cuales se hace cada una de las predicciones puede afectar a la capacidad predictiva. Los tamaños evaluados son $m=5$, $m=10$, $m=15$ y $m=15$. Además de esto se ha estudiado el efecto que tiene añadir una tasa de caída, que consiste en que en cada paso del entrenamiento, las neuronas de entrada (los precios que sirven como input para hacer la predicción) y las neuronas que están en la capa intermedia (número de celdas LSTM como las vistas en el apartado de la estructura de las celdas LSTM, las cuales se ponen en serie) son ignoradas temporalmente con una probabilidad p que se puede definir. Con esto se consigue reducir el sobreajustamiento y que por lo tanto que el modelo generalice mejor. Se ha ajustado por lo tanto la red LSTM en el reduced training set y se ha evaluado en el validation set para cada una de las combinaciones. El resultado se muestra en la tabla 16.

Esta tabla se ha creado mediante la programación de un algoritmo que ha adaptado el problema de predicción a cada una de las situaciones, el cual tiene las siguientes fases:

- Primero se define una m y una tasa de caída prop.
- A partir de la m se crea la tabla 12 en la cual se pasaba la serie temporal a un formato supervisado en el que la última columna era el precio a predecir y el resto, los precios que se daban como input. Esta tabla tiene por lo tanto $m+1$ columnas. De todas las filas de esta tabla se eliminan aquellas que tienen NA, que son las m primeras, ya que si se quiere predecir el precio del día 01/01/2017, que es el primero, en base a los m anteriores, como no se conocen, aparecen como NA hasta que se llega al día $m+1$ para el cual se dispone de los datos de todos los días anteriores. De esta forma lo que se consigue es que en el training no haya ni un solo dato faltante.
- A continuación se lleva a cabo la normalización por secuencia, quedando un resultado como el de la tabla 13.
- Una vez se ha llevado a cabo la normalización se divide toda la tabla en tres partes:
 - reduced training set, que comprende los datos que van del $m+1/01/2017$ hasta el $19/10/2018$, representando el 60% del total de días de los que se dispone.
 - validation set, que comprende los datos que van del $20/10/2018$ hasta el $26/05/2019$, representando el 20% de los días.
 - test set, que comprende el resto de días hasta el $31/12/2019$. Estas tres particiones son las mismas que las usadas en el ajuste del modelo ARIMA, para así, que cuando se hagan predicciones en el test set, estas se puedan comparar.
- Una vez se han definido estas tres particiones, se crea la estructura de la red LSTM, informándole de las dimensiones del input que va a recibir, las cuales van a ser [batch size, time steps, dimensionalidad] que es igual a $[n, m, 1]$. El primer elemento es n , ya que como se ha comentado, se lleva a cabo *mini-batch gradient descent* y por lo tanto los parámetros se van actualizando a partir del promedio de los gradientes de n individuos. El siguiente

elemento es m ya que es el número de time steps que se han definido y finalmente la dimensionalidad es 1 debido a que solo se está utilizando una variable que es el precio de cierre. Se inician los pesos mediante la inicialización de Glorot también llamada inicialización de Xavier (Glorot and Bengio 2010), la cual permite que los gradientes no se desvanezcan o crezcan desmesuradamente y que por lo tanto hace capaz que la red LSTM aprenda las dependencias temporales.

- Se transforma el training set para que tenga tres dimensiones, que es el formato requerido, el cual tiene la siguiente forma: $[642, m, 1]$, siendo 642 el número de elementos del training set, m , el número de time steps y 1 lo que define el número de variables que se usan que en este caso es una, ya que solo se usa el precio de cierre.
- Se ajusta la red LSTM, poniendo un límite de 75 iteraciones o *epochs* (una iteración es actualizar los parámetros 642 veces, tantas veces como días tiene el reduced training set), teniendo como criterio de minimización el error cuadrático medio. Se establece una paciencia de 4, esto es que, si se llega a 4 iteraciones seguidas en las que no se ha mejorado el error cuadrático medio en el validation set se para el entrenamiento, ya que se considera que a partir de ese momento se están sobreajustando los datos y por eso el ECM ya no mejora. Además, una vez ha parado, el modelo acaba teniendo como valor de parámetros aquellos que han dado el mínimo error cuadrático medio. Si el número de iteraciones es $k < 50$, entonces son los parámetros de la iteración $k-3$. Si se llega a 50 iteraciones entonces la iteración de la cual se obtienen los parámetros es una comprendida en el rango $[47, 50]$. El ajuste se ha hecho con el optimizador (Kingma and Ba 2014).
- Una vez se ha ajustado el modelo se asignan sus pesos a otro modelo con la misma estructura que el anterior. Esto se hace debido a que a este modelo se le darán como input los individuos de uno en uno y no de veinte en veinte debido a que solo se quiere llevar a cabo la predicción del día siguiente. Si se intenta hacer las predicciones con el anterior modelo da error debido a que tiene interiorizada el batch size de 20 y no acepta un input con otra dimensión.
- Una vez se ha ajustado el modelo entonces se llevan a cabo las predicciones en el validation set una a una, y cada una de estas predicciones es transformada para deshacer la normalización. Como los datos se han normalizados, al ajustar el modelo, y tal y como se ha comentado en el capítulo de preparación de los datos, lo que se predice no es el precio del día $m+1$, sino la variación del día $m+1$ respecto al primer día de la fila. Es por esto que si se quiere el error cuadrático medio de las predicciones para poder compararlo con el del modelo ARIMA se ha de deshacer esta transformación.
- Finalmente, con las 219 predicciones se calcula el error cuadrático medio y de este se hace la raíz.

Esto se ha hecho para cada una de las combinaciones. El tiempo total que ha llevado la creación de la tabla 16, en la que se muestran los resultados, ha sido de una hora y cuarto pese a solamente utilizar una variable.

Criptomoneda	X	Bitcoin	Ethereum	Eth_classic	Litecoin	Monero	Ripple
m=5	Prop=0	183.3684	7.6064	0.2966	3.0077	3.2247	0.0168
	Prop=0.25	185.8778	7.8863	0.3024	3.1190	3.3074	0.0173
m=10	Prop=0	182.1198	7.4483	0.3034	2.9736	3.1910	0.0172
	Prop=0.25	184.8163	7.6870	0.3014	3.0261	3.2914	0.0177
m=15	Prop=0	186.2850	7.5048	0.2993	3.1038	3.2437	0.0176
	Prop=0.25	183.0627	7.7601	0.2956	2.9963	3.3272	0.0180
m=25	Prop=0	194.7516	8.1913	0.3443	3.2177	3.4782	0.0174
	Prop=0.25	190.5210	7.6289	0.3056	3.0416	3.3694	0.0180

Tabla 16: Tabla en la que se han evaluado la raíz del error cuadrático medio en la predicción de los precios del validation set con diferentes combinaciones de número de time steps y tasas de abandono.

En la tabla 16 se pueden observar varias cosas. La primera es que no hay ninguna criptomoneda para la cual se obtenga el menor error cuadrático medio con una $m=5$. La segunda observación es que la única criptomoneda en la que con una proporción de tasa de caída de 0.25 se ha obtenido un menor ECM es con Ethereum Classic. Las m que han permitido una mejor predicción, las cuales están subrayadas en negro, han sido de 10 para Bitcoin, Ethereum, Litecoin y Monero, de 15 para Ethereum Classic y de 5 para Ripple. Es quizá por el hecho de que Ethereum Classic es criptomoneda con mayor valor de m que con la ayuda de la tasa de caída del 0.25 se consiguen mejores resultados.

La combinación de m y p que se usará para ajustar el modelo a todo el training set y con ello hacer las predicciones en el test set está en la tabla 17. Estas combinaciones son las que minimizan los ECM en el validation set.

Criptomoneda	Bitcoin	Ethereum	Eth_classic	Litecoin	Monero	Ripple
m	10.00	15.00	5.00	5.00	5.00	5.00
prop	0.00	0.00	0.00	0.00	0.00	0.25

Tabla 17: Combinación de m y p utilizadas finalmente

20 Ajuste de todos los modelos en el training set

Una vez ya se conocen las mejores combinaciones de hiperparámetros para cada una de las criptomonedas, tanto en el modelo ARIMA y en las redes LSTM, se procede a ajustar cada uno de los modelos en el training set, que comprende los datos que van del 01/01/2017 al 26/05/2019, esto suponen 866 días, que son el 80% de los datos. Una vez se haya ajustado cada uno de los modelos se llevan a cabo las predicciones de los 219 días restantes, en lo que se llama como test set.

Respecto a la red LSTM, la única diferencia respecto a lo explicado en el apartado de ajuste de la red LSTM es que se han fijado un total de 75 epochs, y no ha habido ningún criterio de parada. Se ha visto que al rededor de este número de iteraciones el error en el training set empezaba

a estabilizarse en todas las criptomonedas y por lo tanto entrenar durante más epochs hubiera supuesto un sobreajustamiento. No se ha mirado el error en el test set para decidir el número de iteraciones, ya que esto hubiera supuesto hacer trampas debido a que en una situación real no se conocen los precios del test que se intenta predecir.

Cabe recalcar que estos últimos 219 individuos (días comprendidos entre 27/05/2019 y el 31/12/2019) no se han usado hasta ahora porque es donde se evaluará la capacidad predictiva final. Si estos días se hubieran tenido en cuenta en el ajuste o a la hora de encontrar los mejores valores de los hiperparámetros entonces los resultados que se hubieran obtenido hubiesen sobreestimado la capacidad predictiva debido a que estarían el modelo hubiera aprendido de los mismos. Este es algo que aunque parezca obvio, a lo largo del trabajo de investigación se ha visto como algunos papers no lo tienen en consideración y lo que hacen es estimar el modelo con todos los datos y a continuación hacer predicciones en una porción de los mismos, obviando el hecho de que todo resultado de dicho estudio va a estar sesgado de una manera importante.

En la tabla 19 se ven los resultados finales de ECM en el test set. Como se puede ver la red LSTM ha obtenido mejores resultados en Ripple y Ethereum mientras que el modelo ARIMA ha obtenido mejores resultados para el resto de criptomonedas. Como ya se ha comentado, el valor nominal del error es diferente según la criptomoneda ya que algunas tienen precios muchos más elevados que otras.

Criptomoneda	ARIMA	LSTM
Bitcoin	372.9561	371.1975
Ethereum	9.8943	8.8398
Eth_classic	0.2744	0.2707
Litecoin	3.8207	3.9050
Monero	3.3990	3.3131
Ripple	0.0131	0.0120

Tabla 18: ECM evaluado en el test set

En los modelos predictivos solo se a utilizado el precio de cierre de periodos anteriores para hacer predicciones sobre el mismo. Hay otros artículos académicos como (Ji, Kim, and Im 2019) en el que se usan casi todas las variables de la blockchain de Bitcoin y en el que se obtienen muy buenas predicciones. Entre estas variables cabe destacar el volumen de transacciones, las tasas aplicadas a las transacciones, los ingresos de los mineros y el tamaño de la blockchain, las cuales son las más correlacionadas. Esto es algo que cabría estudiar hacer para todas las criptomonedas, ya que seguramente con esto, las redes LSTM hubieran obtenido aún mejores resultados que el modelo ARIMA, pero es algo que computacionalmente el equipo usado en este trabajo no soportaba. Pese a ello, mediante el uso de solamente una variable se ha podido hacer una comparación adecuada de la capacidad predictiva de ambos modelos, ya que en el caso contrario, si la red LSTM hubiera dispuesto de más, el modelo ARIMA hubiera estado en desventaja y las conclusiones que se hubieran podido sacar podrían haber estado más motivadas por el hecho de que las primeras disponen de más información más que otra cosa.

En las figuras 37 a 42 se muestran las predicciones hechas con los dos modelos en el test set además del valor real del precio de cierre que está coloreado en negro.

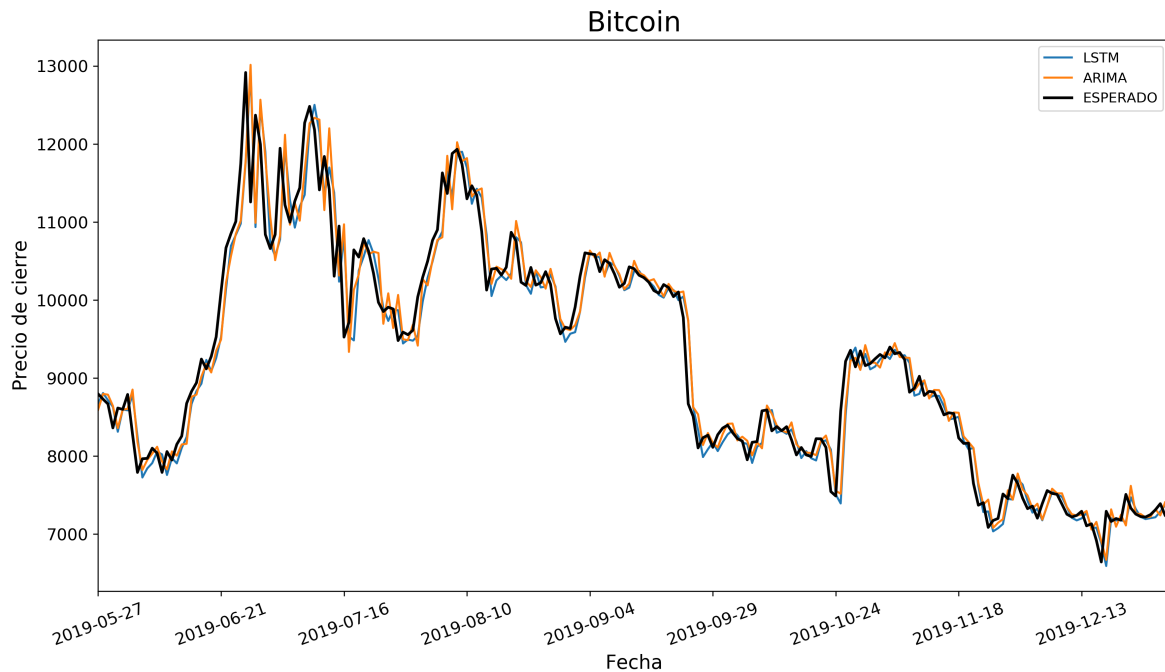


Fig. 37: Predicciones del precio de cierre de Bitcoin en el test set

En la figura 37 se ven las predicciones para Bitcoin. Como se puede observar no hay grandes diferencias respecto a las predicciones de ambos modelos ya que ambos están bastante alineados con lo esperado, que está marcado en negro. Pese a ello, se ve como en el periodo comprendido entre junio y agosto de 2019, la red LSTM parece predecir mejor los cambios de tendencia repentinos.

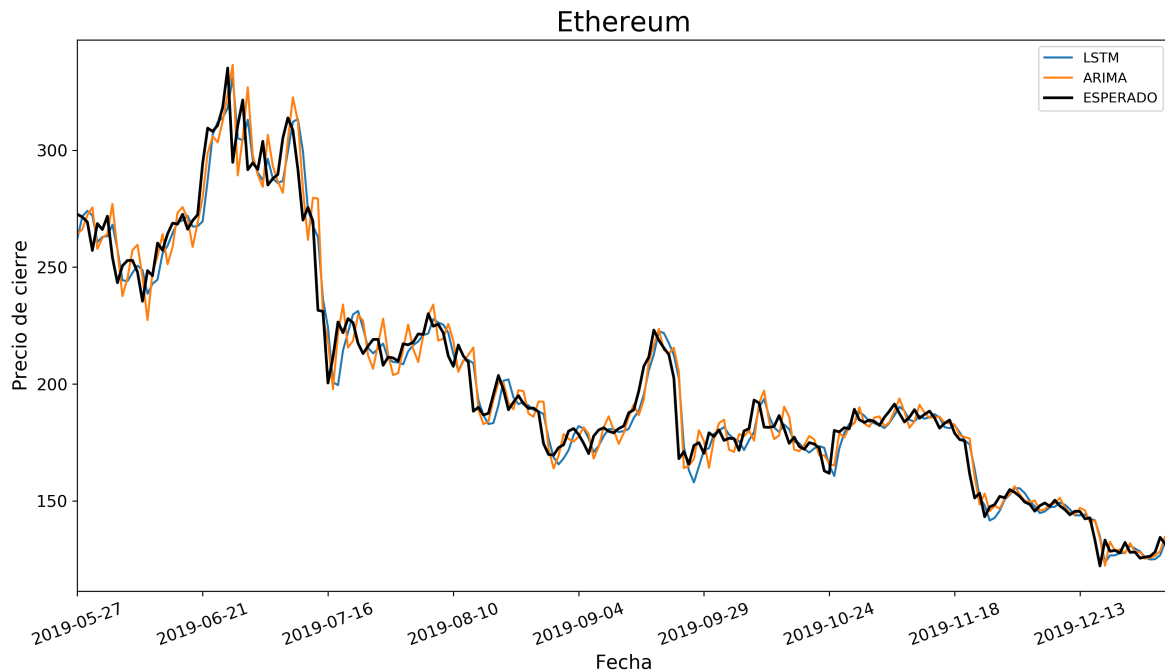


Fig. 38: Predicciones del precio de cierre de Ethereum en el test set

Para el caso de Etheruem presente en la figura 38 se observa un comportamiento parecido al de Bitcoin.

Finalmente en las cuatrocriptomonedas restantes la capacidad predictiva de la red LSTM y el modelo ARIMA las conclusiones son parecidas a las de estos dos gráficos. Cabe destacar que en el caso de Ripple, el modelo ARIMA ha tenido problemas para predecir los cambios de tendencia mientras que la red LSTM si que ha sido capaz de hacerlo.

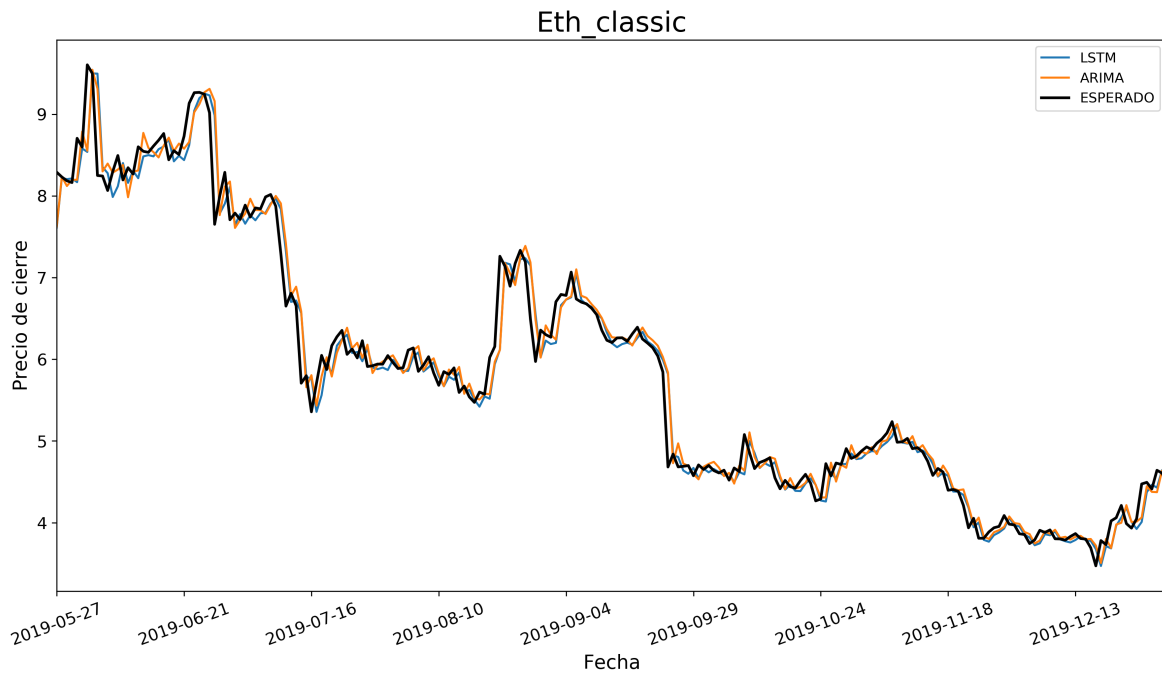


Fig. 39: Predicciones del precio de cierre de Ethereum classic en el test set

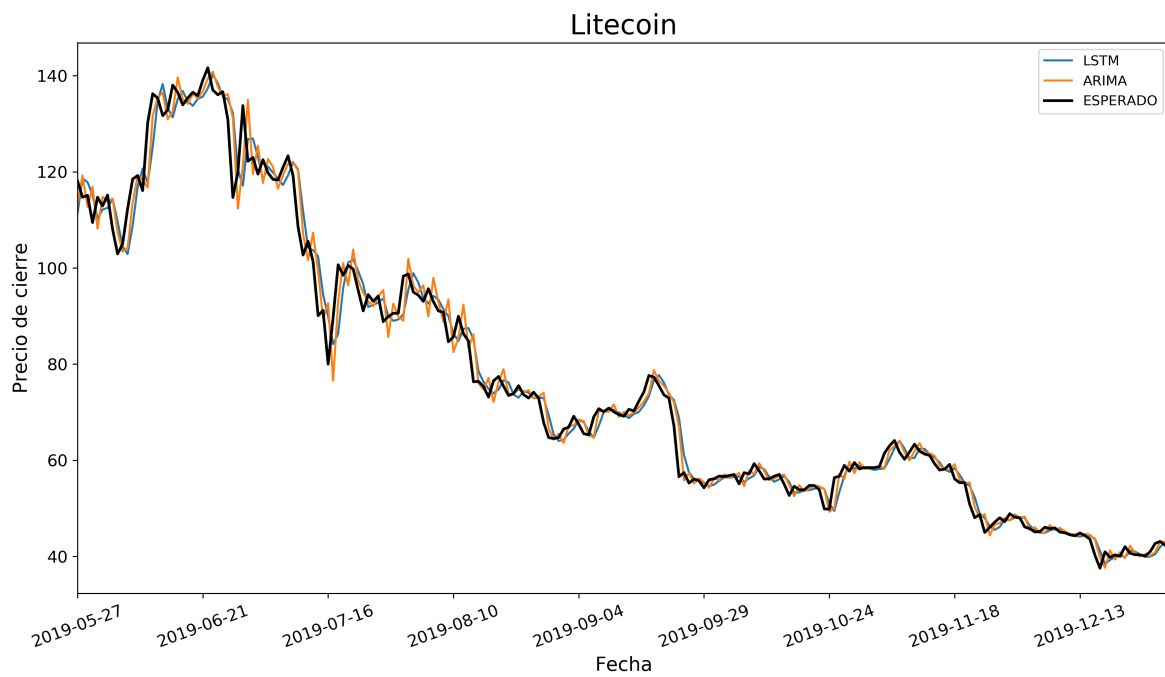


Fig. 40: Predicciones del precio de cierre de Litecoin en el test set

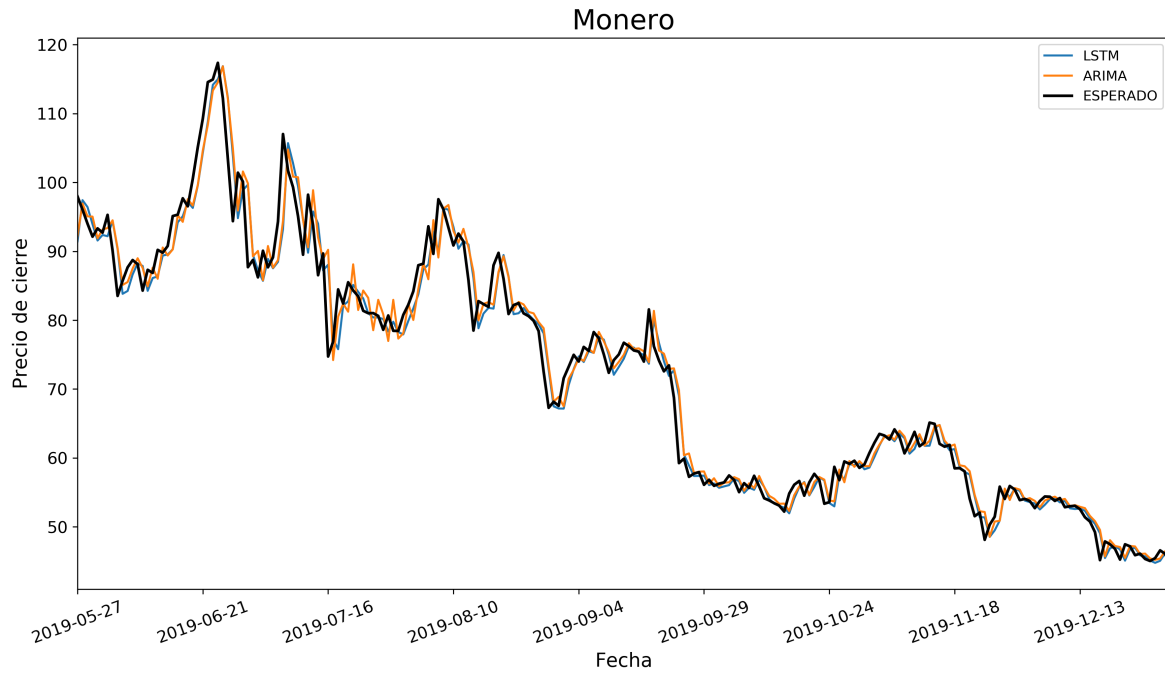


Fig. 41: Predicciones del precio de cierre de Monero en el test set

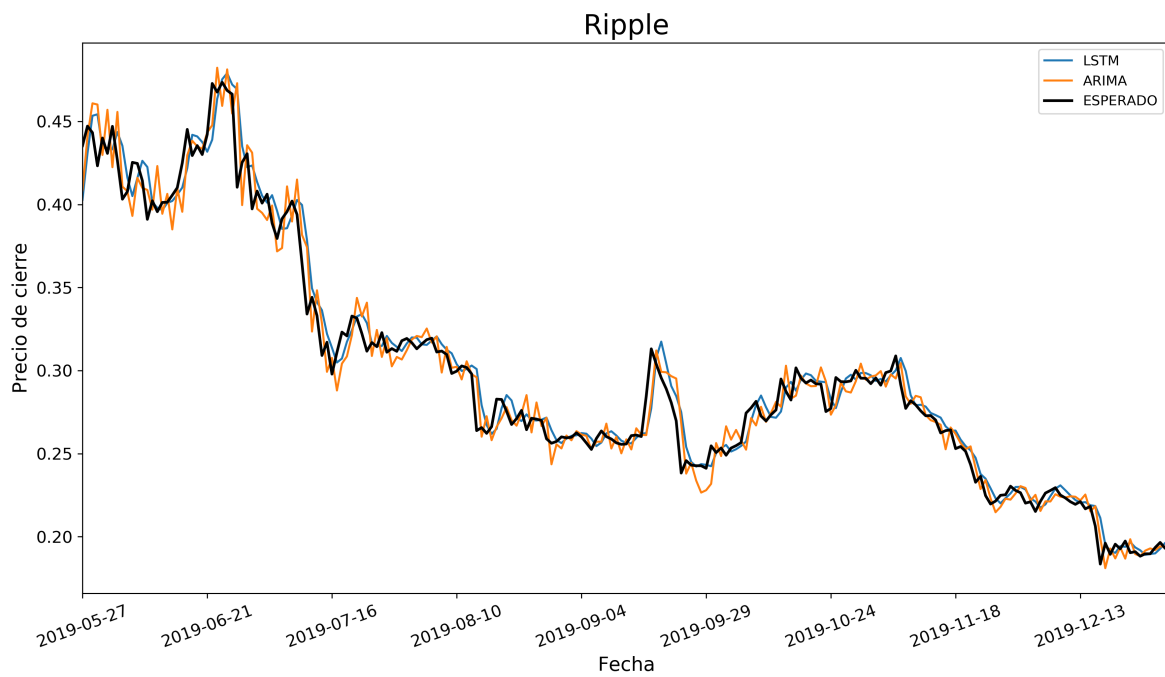


Fig. 42: Predicciones del precio de cierre de Ripple en el test set

20.1 Predicción evolución del precio de cierre

En este apartado se analiza el porcentaje de predicciones en las que se ha predicho correctamente si el precio era superior o inferior al anterior. Esto se ha hecho convirtiendo la variable que contiene las predicciones en una variable binaria, teniendo un valor de 1 si es superior o igual al precio de cierre del día anterior y 0 si es inferior. **No se han ajustado nuevos modelos donde la variable dependiente es binaria, sino que se ha hecho a partir de las predicciones realizadas por los modelos presentados en los apartados anteriores.**

Para tener un referente, en (Ji, Kim, and Im 2019) se utilizaron varios modelos de Machine Learning para la predicción del precio de cierre de Bitcoin a partir de 18 variables, muchas de ellas relacionadas con la blockchain de Bitcoin. Entre los algoritmos usados se encuentran las redes LSTM y el porcentaje de aciertos en la mayoría de los casos está entre un 50% y un 53%. En conclusión, a la hora de predecir como va a evolucionar el precio se esperaban resultados parecidos a los de este trabajo.

Los resultados son los mostrados en la tabla 20. Con el modelo ARIMA, en todas las criptomonedas a excepción de Monero se ha predicho correctamente en menos en un 50% de los casos, por lo tanto, el modelo ARIMA no parece ser adecuado para la predicción de cual va a ser la evolución del precio del día siguiente. Contrariamente a esto, con la red LSTM en todas las criptomonedas todos los porcentajes están por encima del 50%, siendo del 57% para Monero y un 57.53% para Ethereum Classic. En el resto de criptomonedas los porcentajes son próximos al 52% menos para Litecoin, criptomoneda con el que se ha predicho peor.

Estos resultados son mejores de lo esperados con la red LSTM ya que solo se ha utilizado una variable (el precio de cierre) a diferencia de lo hecho en (Ji, Kim, and Im 2019), donde además de utilizar el precio de cierre utilizan 18, algunas de las cuales han sido comentadas en el apartado: ajuste de todos los modelos en el training set

Criptomoneda	ARIMA	LSTM
Bitcoin	49.7717	52.5114
Ethereum	44.2922	52.0548
Eth_classic	45.2055	57.0776
Litecoin	49.3151	48.8584
Monero	50.6849	57.5342
Ripple	47.9452	52.9680

Tabla 19: Porcentaje de aciertos evaluando en el test set

En la figura 43 se muestran los beneficios extraídos de la siguiente estrategia:

- si se predice que el precio del día siguiente es superior al actual, se compra una criptomoneda al precio de cierre actual en caso de no haberlo comprado anteriormente. En caso de acertar, se suma la diferencia del precio a los beneficios. En caso de haberla comprado suponiendo que el precio subía cuando ha acabado bajando, se resta la diferencia de precios a los beneficios.
- si se predice que el precio del día siguiente va a ser inferior, se vende en corto para así

aprovechar la bajada del precio. Se parte del supuesto de que la venta en corto tiene una duración de un día. En caso de no acertar, se resta la diferencia de precios, ya que si se había vendido en corto, el día siguiente se compra la criptomoneda a un precio superior al que se había vendido.

En el caso de la red LSTM, solo en tres de las seis criptomonedas se han obtenido beneficios con esta estrategia, y estas son Ethereum, Ethereum Classic y Monero. Cabe destacar que mientras para las dos primeras los beneficios se mantienen positivos a lo largo de todo el periodo, con Monero, pese a que es la criptomoneda con la que se ha obtenido mejor % de predicciones correctas, hasta agosto de 2019 se mantienen negativos. Para las otras tres, con Litecoin y Ripple los beneficios tienen una tendencia a la baja constante, mientras que con Bitcoin se llega a tener una pérdida de 6000 dólares pero se acaba recuperando las pérdidas, estando cerca de estar como al inicio al final del periodo.

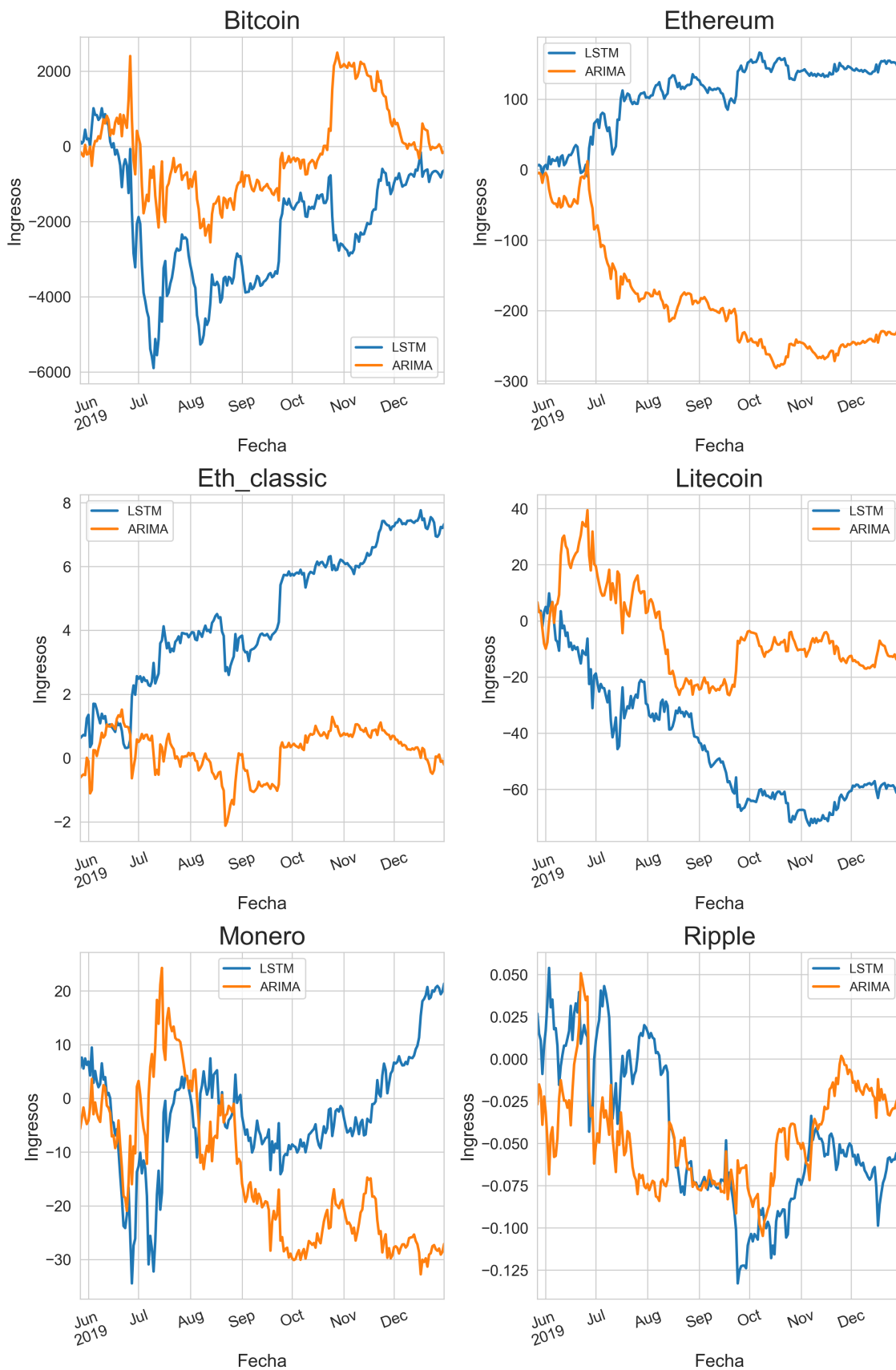


Fig. 43: Predicciones del precio de cierre de Ripple en el test set.

A diferencia de la red LSTM, con el modelo ARIMA no se acaba con beneficios positivos en ninguna criptomoneda al final del periodo, pero cabe destacar que con Bitcoin, Litecoin y Ripple, criptomonedas en las que la red LSTM ha fracasado en tener beneficios positivos, se acaba con un mejor resultado. En Litecoin esto puede ser debido a que se trata de la única criptomoneda en la que el modelo ARIMA ha predicho mejor que la red LSTM la evolución del precio, tal y como se ha visto en la tabla 19. Para Bitcoin y Litecoin se ha investigado el porqué de esto y se ha llegado a la conclusión de que es debido a que la red LSTM ha predicho peor que el modelo ARIMA en el mes de junio, periodo en el que se producen las mayores variaciones de precios. En terminos generales, pese a que un algoritmo tenga mayor porcentaje de aciertos que otro, si el que tiene menos predice mejor la evolución en los momentos de más volatilidad puede pasar algo como lo que se ve en esta figura.

Finalmente, en la figura 44, se muestra lo mismo que en la anterior, pero los beneficios están divididos entre el precio medio de cada criptomoneda en este periodo, para de esta forma tener un beneficio relativo al precio de cierre. Esto se hace debido a que Bitcoin tiene un precio miles de euros superior al resto de criptomonedas y por lo tanto hace que con la figura anterior no se puedan llevar a cabo comparaciones entre las criptomonedas.

En el caso de la red LSTM, como se puede observar, con Ethereum se acaba con un beneficio que es un 125% superior al precio de cierre medio de la misma en este periodo, seguido de Ethereum con un valor cercano al 75% y de Monero con un 30%. El peor resultado es con Litecoin, con el que se tienen unas pérdidas del 80%.

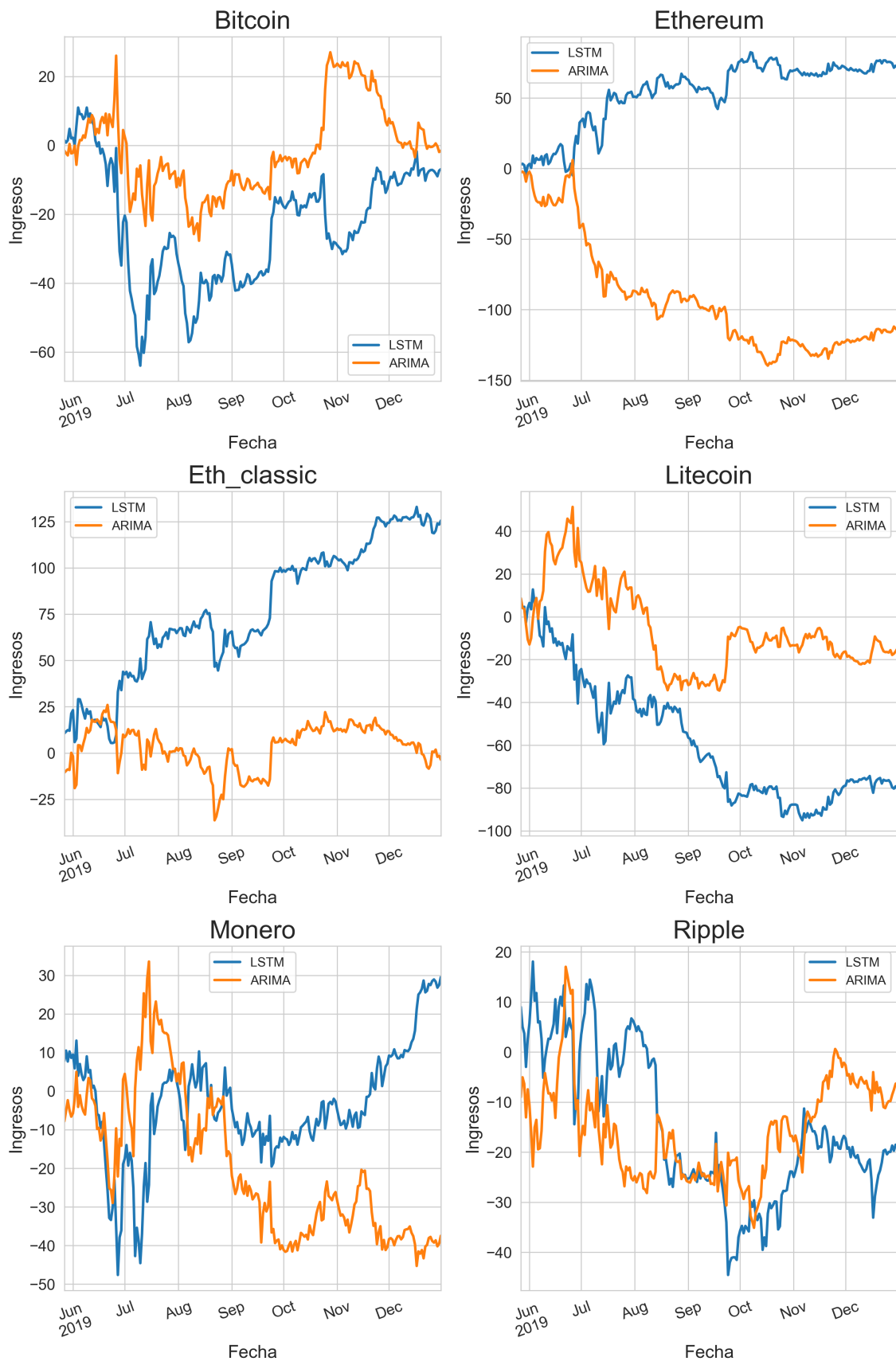


Fig. 44: Predicciones del precio de cierre de Ripple en el test set.

De estas tres criptomonedas con las que se ha obtenido beneficios, dos de ellas, que son Ethereum Classic y Monero se encuentran entre las criptomonedas con mayor riesgo atendiendo al criterio de VaR y CVaR estudiado en el apartado de “Análisis estadístico de las criptomonedas”, siendo Ethereum Classic la que tiene más. Por otro lado, Ethereum se encontraba en un punto medio entre las criptomonedas con mayor y menor riesgo.

Por lo tanto, por un lado está Ethereum Classic, criptomoneda con la que se han obtenido mayores beneficios relativos pero a la vez la que tiene asociado mayores riesgos y por otro está Ethereum, que tiene menor riesgo que las otras dos pero también reporta menos beneficios. En conclusión, la decisión de en cual se debería invertir dependería de la propensión al riesgo de la persona que invierte.

Es necesario destacar que todo esto se ha de tomar todo con mucha cautela, ya que el periodo estudiado es solo de tres años y por lo tanto hacen falta más años de datos históricos para sacar conclusiones más fiables. Además, solo se ha utilizado una variable, que es el precio de cierre, debido a la capacidad computacional del equipo utilizado en el trabajo. Se debería estudiar, como cambia esta situación con el uso de más variables, como por ejemplo las relativas a la blockchain de cada una de estas criptomonedas además del precio máximo, precio mínimo, o precio de apertura.

21 Conclusiones

Desde la aparición de Bitcoin ha habido una constante creación de criptomonedas, llegando a haber más de 200 en la actualidad. A medida que pasan los años, su influencia cada vez es más evidente. Tienen una serie de características que las hacen únicas, como las de permitir realizar pagos sin intermediario, de forma muy rápida, sin apenas costes, junto a su nivel de descentralización y el uso de la tecnología blockchain, y finalmente el ser un medio muy seguro y que puede ser utilizado a nivel mundial por cualquier persona con acceso a internet. Todo esto ha hecho que sean foco de análisis, imitación e inspiración.

Ha habido mucha discusión sobre la posibilidad de coexistencia de las criptomonedas con el dinero fiat o sobre la posibilidad de que en un futuro haya un criptosistema. A lo largo del trabajo se ha acabado concluyendo que esto es una posibilidad muy remota. La mayoría de las criptomonedas tienen un problema respecto al dinero fiat y es la endogeneidad de su precio, el cual viene determinado por el uso que se le da y el valor que la sociedad considera que tiene, ya que estas, como se ha visto, no tienen valor intrínseco. Además, debido a que las transacciones, en criptomonedas como Bitcoin, tienen un límite de espacio, en caso de pánico, puede haber una alta competencia entre poseedores de bitcoins, que puede provocar una alza muy grande de las tarifas, lo que hace que su uso transaccional se vea perjudicado en caso de que se dé un alto uso.

Esta inestabilidad financiera de un criptosistema también puede venir originada por otras razones como son la estabilidad de la que presumen debido a que, según ellos, un criptosistema no sufriría de inflación al estar limitada la oferta de criptomonedas (si estuviera basado en una criptomoneda como Bitcoin). Esto puede hacer que los agentes económicos tomen más riesgos y que por lo tanto, ante un pequeño evento, las consecuencias económicas sean mucho más graves que si se estuviera en una situación en la que la población tomara más precauciones.

Si actualmente se produce una crisis, el banco central tiene la capacidad de llevar a cabo una expansión cuantitativa u operaciones de mercado abierto, pero en un criptosistema, en caso de no disponer de estas herramientas, los efectos de la crisis serían mucho más graves, ya que no se podría dar respuesta a estas situaciones.

Otro problema es el relacionado con la creación de M1, M2 y M3 cripto, las cuales serían reclamaciones de criptos y no estarían registradas en ninguna blockchain, por lo tanto, las criptomonedas perderían su rasgo distintivo de tener todas las transacciones registradas en la blockchain y se convertirían en una alternativa muy parecida al dinero fiat de curso legal. La gran diferencia está en que al ser reclamos sobre monedas que no están registrados en ninguna blockchain, en caso de temor, la población empezaría a preocuparse de una forma desmesurada sobre si las cripto M1, M2 y M3 que tienen son equivalentes a las criptomonedas M0 que existen, haciendo que el criptosistema fuera muy inestable.

Aunque el dinero fiat es indiscutiblemente el tipo de dinero con más presencia en las centenas de millones de transacciones nacionales e internacionales que se producen a diario a nivel mundial, hay una serie de problemas asociados a las mismas. Como su uso tiende a requerir del acceso a servicios financieros, proporcionados generalmente por entidades bancarias u operadores de transferencia dinero, en países en desarrollo como los pertenecientes a África subsahariana, su uso se ha visto limitado por el bajo nivel de bancarización de los mismos. Por otro lado, si a esto

se junta el alto coste de enviar remesas a nivel internacional, sobre todo en esta región, aparece la necesidad de otras alternativas.

Aunque es una realidad que se ha visto reflejada en el análisis estadístico que hay una tendencia a la baja de los costes de enviar remesas a nivel internacional, debido sobre todo a la presión del G20 y ONU, siguen habiendo dos problemas que aún no se han solucionado: el primero es que la banca, en estas operaciones, sigue teniendo unos costes muy por encima de su máximo competidor y medio más utilizado, que son los Operadores de Transferencia de Dinero (OTD). En caso de enviar 200 dólares en remesas, en el año 2019, los costes promedios asociados son del 10.36% y 6.02% respectivamente, haciendo evidente porque los segundos son los más utilizados. El segundo problema surge cuando se pone el foco en África subsahariana, ya que en esta región los costes promedio son bastante más elevados que para el resto de regiones del mundo, siendo un 50% inferiores en caso de usar un OTD respecto a la Banca. Pese a esto, al igual que la tendencia general, los costes promedio en África subsahariana se están reduciendo, pasando de un 9.67% a un 7.86% del año 2011 al 2019 para el caso de los OTD.

Aunque este es uno de los motivos de la aparición de iniciativas como Libra, hay otra razón más importante detrás de esto, y es el nivel de bancarización de los países menos desarrollados del mundo. A partir de la base de datos de Findex se han sacado muchas conclusiones respecto cual es el estado actual y la evolución en los últimos años de variables relacionadas con esto. Se ha observado como el nivel de bancarización, medido como adultos (%) con una cuenta en un banco o que han utilizado un servicio monetario mediante un móvil, está en niveles cercanos al 40% en regiones como África subsahariana, muy lejano del 91.67% de las economías de altos ingresos. Pese a esto, desde el 2011, año en el que en África subsahariana estaba en el 20%, se ha doblado este indicador. El resto de agrupaciones de países también han tenido una tendencia alcista pero siguen estando por debajo del 60%.

Algo en lo que sí destaca África subsahariana es en los altos niveles de adultos (%) que usan un servicio monetario con un teléfono o que han enviado o recibido remesas a nivel regional a través del teléfono móvil, variables en las que África subsahariana está en un 24%, muy por encima del resto mundo. Esto, como se ha visto, es debido a varios factores entre los que se encuentra la falta de infraestructuras en estos países o la imposibilidad de tener acceso a la banca por parte de la mayoría de la población. La causa que hay detrás son otras alternativas como son M-Pesa, servicio proporcionado por Vodafone desde 2009 dentro de las fronteras de algunos países de África subsahariana que ha permitido que decenas de millones de habitantes tengan acceso a servicios monetarios por primera vez. Todo esto gracias a que las transferencias son realizadas a través del teléfono móvil de una forma tan sencilla como enviar un SMS. Esto no hubiera sido posible si no fuera por el hecho de que en todas las agrupaciones de países estudiadas, incluido África subsahariana, tienen más de un 60% de los adultos con teléfono móvil. Cabe destacar que con los proyectos que pretenden hacer llegar el internet a todo el territorio africano, como por ejemplo el cable de fibra óptica submarino que Facebook quiere construir, esta situación puede cambiar drásticamente en el futuro, pudiendo tener implicaciones muy importantes sobre el uso de tecnologías como la de Libra de una manera más generalizada en dicho territorio.

Aunque M-Pesa está teniendo una incidencia muy importante en el acceso a servicios financieros de las regiones más desfavorecidas, este problema continúa persistiendo, ya que aunque en países como Kenia, donde M-pesa lleva más años trabajando, hay más de un 80% de los adultos con

acceso a una cuenta y más de un 70% que utilizan un servicio monetario con un teléfono móvil, en el resto de países de África subsahariana se está muy lejos de estos números y por lo tanto aún queda mucho por hacer. Además, en regiones como India, Sud África, Albania o Rumanía M-Pesa ha sido un absoluto fracaso y ha tenido que abandonarlos. Es por esto que M-Pesa es necesaria pero insuficiente. Ésta ha servido para que en los países de África subsahariana en los que ha estado haya un nivel de accesos servicios financieros superior a los que no han tenido este servicio, pero aún falta mucho por hacer.

Por razones como el bajo nivel de bancarización de algunas zonas del mundo, y debido al alto coste de enviar remesas a nivel internacional, el 18 de junio de 2019 se anunció Libra. Pese a la insistencia en que se trata de una criptomoneda por numerosos medios, al basarse en el uso de una blockchain bastante centralizada y permitida y al estar creada por una corporación tecnológica como Facebook, entra dentro de lo que se denomina monedas digitales corporativas. Nació con el objetivo de dar respuesta al hecho de que en 2018 había 1700 millones de adultos fuera del sistema financiero, 1000 millones de los cuales tienen acceso a teléfono celular. A diferencia de las criptomonedas, las cuales tienen limitaciones bastante importantes respecto al número de transacciones que pueden ser registradas en la blockchain, la blockchain privada de Libra, al estar administrada por los miembros de la Asociación Libra, entidades que han pagado un mínimo de 10 millones para formar parte de ella, y al no ser necesaria la actividad del mining, se puedan procesar transacciones a un ritmo mucho más elevado que con las criptomonedas, ya que no se requiere la resolución de ningún problema computacional.

Esta moneda digital, al tener un valor que depende de una cesta de monedas como son el dólar y el euro, y al poseer una reserva de activos que hace que tenga valor intrínseco, evitan que el valor de Libra dependa del estado del mercado y además la convierten en una moneda muy estable, todo lo contrario a las criptomonedas que no son stablecoins. Ha sido por el potencial que tenía para convertirse una moneda que podía llegar a poner en peligro la estabilidad financiera que Facebook ha tenido que ceder y ahora el proyecto soportará transacciones con Libra y las llamadas single-currency stablecoins, las cuales representan monedas nacionales como el dólar o el euro.

Ahora hace falta ver con estos cambios cuantos países van a aceptar el servicio Libra, sobre todo aquellos países con monedas poco estables donde Libra podría convertirse en una moneda refugio en presencia de inestabilidad, provocando en consecuencia la depreciación de la moneda local. Pese a esto, cabe destacar que como ahora existirán otras alternativas como las SCS esto tendrá dos consecuencias: la primera es que aquellos que desconfían en Libra tienen una opción muy segura como son las SCS y en segundo lugar, los bancos centrales, que estaban preocupados por el poco control que iban a tener de Libra, ahora dispondrán de más, ya que muchas de sus transacciones serán realizadas con SCS, las cuales, para crearse, requieren de un respaldo 1:1, y por lo tanto, no habría creación neta de dinero, aspecto que habría causado preocupación.

Otro aspecto que cabría estudiar es que razones hay detrás del fracaso de M-Pesa en regiones como India, que tiene 1353 millones de habitantes en 2018. Tanto Libra como M-Pesa son servicios pensados para poder realizarse a través del móvil, y en el caso de Libra quiere postularse como una alternativa a los medios convencionales para transferir dinero a nivel internacional, a un bajo coste. Pese a que se trata de un proyecto que puede ser revolucionario, el uso por parte de la sociedad, la aceptación por parte de los gobiernos y bancos centrales, y la capacidad que va a

tener para cumplir con todas las regulaciones de los países en los que quiera establecerse todavía es una incógnita.

Pese que el uso del dinero digital aún es algo que está lejos de aplicarse a nivel mundial, en países como Suecia solo uno de cada ocho suecos realizan compras en efectivo hoy en día debido a que usan una aplicación móvil para realizar las transferencias.

Los bancos centrales de algunos países como China o Canadá, paralelamente a todo lo que está pasando, están estudiando la creación de monedas digitales administradas por ellos. Éstas podrían permitir mayor eficiencia en la realización de pagos, más seguridad y más control del estado sobre la población. Como se ha visto, esto podría permitir al banco central usar las CBDC para afectar la oferta y la demanda de una manera inmediata cambiando el valor de la misma, sin necesidad de recurrir a operaciones de mercado abierto.

Podrían estar basadas en cuentas en control del banco central o basadas en tokens, en cuyo caso podría utilizarse una blockchain no permissionada, o permissionada. El problema del primer tipo de blockchain es el asociado a los problemas de escalabilidad, y por lo tanto requeriría una mejora de la tecnología Blockchain. Aunque puede ser algo que los bancos centrales están estudiando, hay aspectos que hacen pensar que no van a llegar a realizarse. En caso de estar basada en la tecnología blockchain, y al convertirse en proveedor de servicios, necesitaría construir una enorme infraestructura para ello. En segundo lugar, al sustituir a los bancos comerciales, sería responsable de dar créditos a la población, algo que no entra dentro de las funciones de un banco central.

En la segunda parte del trabajo se ha realizado primero de todo un análisis estadístico de las criptomonedas. En el primer análisis exploratorio se ha visto como Bitcoin, debido a su enorme popularidad, se trata claramente de la criptomoneda que mayor precio tiene, muy por encima del resto y además es la criptomoneda con menor coeficiente de variación. Por otro lado Ripple muestra las características opuestas, con un precio medio cercano a 0 y un coeficiente de variación elevado.

La rentabilidad continua de las criptomonedas presenta unas características bastante diferentes. En este caso es Ethereum la que tiene menor coeficiente de variación y Ethereum Classic la que más. Todas tienen una rentabilidad continua promedio mayor que 0 y todas, a excepción de Bitcoin, tienen una asimetría positiva.

Mediante el criterio de AIC se ha podido concluir primero de todo, que la distribución normal no se encuentra entre las distribuciones que mejor ajustan los datos de la rentabilidad continua y en segundo lugar que la distribución normal inversa gaussiana aparenta ser la más adecuada de entre las usadas, aspecto confirmado mediante el test de Kolmogorov, que no ha dado significativo para ninguna de las criptomonedas. Pese a esto, hubiera sido interesante el uso de otras distribuciones como la distribución generalizada t de student o la distribución generalizada hiperbólica, que han dado buenos resultados en otros estudios, pero que no ha sido posible utilizar.

Para las mejores distribuciones se han estimado los parámetros mediante máxima verosimilitud y se ha calculado el riesgo asociado a cada criptomoneda mediante el VaR y el CVaR, concluyendo que Monero, Ethereum Classic y Ripple son las criptomonedas con mayor riesgo asociado y

Bitcoin la que menos tiene.

En la última parte del trabajo se han utilizado dos modelos que son el modelo ARIMA y la red LSTM para predecir el precio de cierre de las criptomonedas con un día de posterioridad. Tal y como se ha mostrado, pese a que los modelos estadísticos poseen una serie de propiedades asociados a los parámetros estimados, estos modelos no garantizan una buena capacidad predictiva, ya que, tal y como se ha demostrado, un modelo sesgado subespecificado puede llegar a hacer mejores predicciones que un modelo perfectamente especificado, siempre y cuando se cumplan una serie de condiciones. Esto realmente no supone un problema debido a que los modelos estadísticos, en términos generales, son modelos explicativos que buscan establecer relaciones entre una variable explicativa y un conjunto de covariables y no tanto llevar a cabo predicciones en nuevos datos. Es en esto último donde destacan los algoritmos de Machine Learning.

Para los modelos ARIMA se ha optado por utilizar el criterio de AIC para escoger aquellos que proporciona un mejor ajuste, y esto se ha hecho para modelos de diferente complejidad en cada una de las criptomonedas para así evitar que se produzca un sobreajustamiento. Este sobreajustamiento se ha podido observar mediante el error cuadrático medio evaluado en el validation set, ya que ha sido en los modelos más complejos en los que las predicciones eran peores atendiendo a la función de pérdida del error cuadrático medio. Con esto se ha demostrado la afirmación del anterior párrafo. Un modelo que ajusta mejor los datos, en este caso, según el criterio de AIC, no tiene porque ser el modelo que mejor predicciones hace.

El otro modelo utilizado han sido las redes LSTM debido a la capacidad que éstas tienen de captar las dependencias temporales sin entrar en problemas del gradiente desvaneciente muy común en la red neuronal recurrente vanilla. La normalización utilizada ha sido la normalización por secuencia por la capacidad que esta tiene de captar las tendencias locales. Para actualizar los parámetros se ha optado por utilizar Adam Optimization con batch size=20, ya que se ha observado un empeoramiento de las predicciones cuando un tamaño de batch size=1 en el entrenamiento. En la elección de la longitud de las secuencias con la que se iba a ajustar cada criptomoneda se ha visto como ésta no ha sido la misma para todas ellas. Bitcoin se ha ajustado con una longitud de secuencia de 10, Ethereum de 15 y el resto de criptomonedas de 5.

Respecto la capacidad predictiva, evaluada mediante el error cuadrático medio cometido por el modelo ARIMA y al red LSTM a la hora de predecir los precios de cierre del test set, la red LSTM ha tenido un error cuadrático medio inferior al modelo ARIMA en todas las criptomonedas menos Ethereum.

Caben destacar varias cosas. La primera es que, a diferencia de otros estudios mencionados en este trabajo, donde utilizan muchas variables relacionadas con cada criptomoneda, en éste se ha utilizado el precio de cierre por las limitaciones que ya han sido explicadas. Por lo tanto, no se ha explotado al máximo el potencial que las redes LSTM, debido a éstas pueden aprender dependencias temporales de muchas variables a la vez y, por lo tanto, esta capacidad predictiva encontrada puede llegar a ser mejorada de forma importante.

Un segundo hecho a destacar es que el horizonte temporal utilizado ha sido de tres años, ya que es lo que hay disponible para la mayoría de estas criptomonedas. De bitcoin, para el que se disponen datos de los últimos 11 años, solo se han utilizado los últimos tres años, así que cabría estudiar

el efecto que tiene ampliar el horizonte. Pese a esto, si el comportamiento de una criptomoneda varía a lo largo de su existencia, sería interesante analizar si realmente sale a cuenta hacer que la red LSTM aprenda dependencias temporales de hace años, cuando se daba un comportamiento que no se da ahora.

El último hecho a destacar es que el periodo final evaluado ha sido de más de 7 meses, pero es insuficiente. Hace falta esperar algunos años para disponer de más datos de las criptomonedas y así poder sacar conclusiones sobre la capacidad predictiva de estos modelos, ya que podría haberse dado el caso de que el periodo evaluado tuviera unas características intrínsecas, como una tendencia bajista constante, que han posibilitado que las predicciones sean más fáciles o más difíciles de hacer, afectando así a las conclusiones realizadas en el trabajo.

Gráficamente no se han observado diferencias importantes entre las predicciones del modelo ARIMA y la red LSTM aunque cuando se ha evaluado la capacidad predictiva de la evolución del precio, ya sea ascendente o descendente, se han observado grandes diferencias. Ha sido en esto donde la red LSTM ha destacado bastante por encima del modelo ARIMA, consiguiendo una **accuracy** superior al 50% en todas las criptomonedas a excepción de Litecoin. En el caso del modelo ARIMA, solo se ha conseguido esto para Monero. Teniendo en cuenta que solo se ha utilizado el precio de cierre y que casi todos los estudios se quedan muy cerca del 50%, los resultados han sido mejores de lo esperados.

Estas predicciones sobre la evolución ascendente o descendente del precio han sido las utilizadas para analizar una estrategia de trading muy sencilla. Ésta solo ha resultado rentable para Ethereum, Ethereum Classic y Monero considerando que no hay costes asociados a las transacciones para así simplificar. Hace falta estudiar como, dados los cambios mencionados anteriormente, como son el uso de más años o más variables, puede hacer cambiar estos resultados, ya que si solo con una variable se han conseguido estos resultados, cabe esperar, seguramente, mejores resultados si se usan variables que puedan explicar de alguna manera la evolución del precio de las criptomonedas.

Lista de figuras

1	Ejemplo de aplicación de la función SHA256 Hash. Fuente: A web based demonstration of blockchain concepts, Github	11
2	Blockchain formada por cuatro bloques en la que se muestran los dos últimos. Contiene un conjunto de transacciones inventadas que sirven de ejemplo. Fuente: A web based demonstration of blockchain concepts, Github	12
3	Ejemplo de creación de la llave pública a partir de la llave privada. Fuente: A web based demonstration of blockchain concepts, Github	13
4	Creación de la firma de una transacción a partir del mensaje y la llave pública del que la realiza. Fuente: A web based demonstration of blockchain concepts, Github	14
5	Blockchain formada por cuatro bloques. Fuente: A web based demonstration of blockchain concepts, Github	15
6	Millones de bitcoins existentes desde su creación junto a la progresión que tendrán hasta 2033. Fuente: ¿Son 21 millones de bitcoins suficientes?	26
7	Evolución de M0, M1 y M2 durante la Gran Depresión, 1929-1933. Fuente: (Danielsson 2019)	31
8	Riesgo real y riesgo percibido. Fuente: El riesgo endógeno - Endogenous risk . . .	32
9	Total de remesas (millones de dólares estadounidenses) recibidas (gráficos de arriba) y enviadas (gráficos de abajo) por 7 agrupaciones de países del año 2000 al año 2019. Elaborado a partir de: The World Bank, Remittance Prices Worldwide, disponible en http://remittanceprices.worldbank.org	44
10	Los 10 países con mayor entrada de remesas (millones de dólares estadounidenses) en el año 2019. Elaborado a partir de: The World Bank, Migration and Remittances Data	45
11	Los 10 países con mayor entrada de remesas (millones de dólares estadounidenses) en el año 2018. Elaborado a partir de: The World Bank, Migration and Remittances Data	45
12	Coste promedio de enviar remesas del año 2000 al año 2019 según la agrupación de países que las recibe (gráficos de arriba) o las agrupaciones que las envían (gráficos de abajo). Elaborado a partir de: The World Bank, Remittance Prices Worldwide, disponible en http://remittanceprices.worldbank.org	47
13	Coste promedio de enviar remesas según la institución encargada de hacer la transferencia del año 2011 al 2019. Elaborado a partir de: The World Bank, Remittance Prices Worldwide, disponible en http://remittanceprices.worldbank.org	48
14	Adultos que informan tener una cuenta (por sí mismos o junto con otra persona) en un banco u otro tipo de institución financiera o informan personalmente utilizando un servicio monetario mediante un móvil en los últimos 12 meses para los años 2011, 2014 y 2017. Elaborado a partir de: The Global Findex Database 2017	50

15	Adultos de los países de África subsahariana que informan tener una cuenta (por sí mismos o junto con otra persona) en un banco u otro tipo de institución financiera o informan personalmente utilizando un servicio de dinero móvil en los últimos 12 meses para los años 2011, 2014 y 2017. Se han añadido dos agrupaciones nuevas que son M-Pesa, países que tienen en el servicio de M-Pesa y No M-Pesa, países que no disponen de este servicio. Elaborado a partir de: The Global Findex Database 2017	51
16	Adultos que disponen de teléfono móvil en el año 2017 según agrupación de países a la que pertenecen. Se han añadido las dos agrupaciones extra de M-pesa y No M-Pesa. Elaborado a partir de: The Global Findex Database 2017	52
17	Adultos que reportan haber utilizado un servicio monetario mediante un teléfono móvil para los años 2014 y 2017 según agrupación de países a la que pertenecen. Se han añadido las dos agrupaciones extra de M-pesa y No M-Pesa. Elaborado a partir de: The Global Findex Database 2017	53
18	Adultos que reportan haber utilizado un servicio monetario mediante un teléfono móvil para los años 2014 y 2017 según país de África al que pertenecen. Elaborado a partir de: The Global Findex Database 2017	54
19	Adultos que han enviado o recibido remesas a través del móvil en los años 2014 y 2017 según la agrupación de países a la que pertenecen. Elaborado a partir de: The Global Findex Database 2017	55
20	Adultos que han enviado o recibido remesas a través del móvil en los años 2014 y 2017 según país de África subsahariana al que pertenecen. Elaborado a partir de: The Global Findex Database 2017	56
21	Histogramas de la rentabilidad continua de las seis criptomonedas del 1 de enero de 2017 al 31 de diciembre de 2019	61
22	Histogramas de la rentabilidad continua de las seis criptomonedas junto a la función de distribución de probabilidad que mejor las ajusta y los parámetros de las mismas del 1 de enero de 2017 al 31 de diciembre de 2019	65
23	QQ-Plots de la distribución que mejor ajusta la rentabilidad continua de las seis criptomonedas del 1 de enero de 2017 al 31 de diciembre de 2019	69
24	PP-Plots de la mejor distribución para la rentabilidad continua de las seis criptomonedas del 1 de enero de 2017 al 31 de diciembre de 2019	70
25	VaR para la distribución que mejor ajusta los datos de la rentabilidad continua de las seis criptomonedas del 1 de enero de 2017 al 31 de diciembre de 2019	72
26	CVaR para la distribución que mejor ajusta los datos de la rentabilidad continua de cada una de las seis criptomonedas del 1 de enero de 2017 al 31 de diciembre de 2019	74
27	Serie temporal del precio de cierre para las seis criptomonedas del 1 de enero de 2017 al 31 de diciembre de 2019	84
28	Neurona de RNN sin desarrollar. Fuente: Understanding RNN and LSTM, Aditi Mittal	89
29	Neurona desenrollada a través del tiempo con función de activación tangente hiperbólica. Fuente: (Ji, Kim, and Im 2019)	90
30	Función de activación tangente hiperbólica. Fuente: (Géron 2019)	91

31	Gradiente descendiente aplicado a la función $y=x^2$ con una tasa de aprendizaje baja. El eje horizontal representa los valores que toma x , el eje vertical los valores de $f(x)$. Fuente: Gradiente Descendiente para aprendizaje automático, Jose Martinez Heras	92
32	Gradiente descendiente aplicado a la función $y=x^2$ con una tasa de aprendizaje elevada. El eje horizontal representa los valores que toma x , el eje vertical los valores de $f(x)$. Fuente: Gradiente Descendiente para aprendizaje automático, Jose Martinez Heras	93
33	Ejemplo de retropropagación con $m=3$. Las flechas negras y representan los cálculos que se hacen en la fase de propagación hacia adelante, las flechas rojas los cálculos que se hacen en la retropropagación para actualizar los parámetros	94
34	Gradient Descent aplicado con tres metodologías para la resolución de un problema de regresión con dos parámetros. Fuente: [atgeron2019hands]	96
35	Celda LSTM. Fuente: (Géron 2019)	99
36	Dos celdas LSTM apiladas para la predicción de un precio en base al precio de los tres días anteriores. Fuente: (Nicholas et al. 2018)	102
37	Predicciones del precio de cierre de Bitcoin en el test set	107
38	Predicciones del precio de cierre de Ethereum en el test set	108
39	Predicciones del precio de cierre de Ethereum classic en el test set	109
40	Predicciones del precio de cierre de Litecoin en el test set	109
41	Predicciones del precio de cierre de Monero en el test set	110
42	Predicciones del precio de cierre de Ripple en el test set	110
43	Predicciones del precio de cierre de Ripple en el test set.	113
44	Predicciones del precio de cierre de Ripple en el test set.	115

Lista de tablas

1	Coste promedio de enviar remesas mediante el uso de un banco o un Operador de Transferencia de Dinero para los años 2011, 2015 y 2019 según la agrupación de países que las recibe. Elaborado a partir de: The World Bank, Remittance Prices Worldwide, disponible en http://remittanceprices.worldbank.org	49
2	Estadísticos de resumen del precio de cierre de Bitcoin, Ethereum, Ethereum Classic, Ripple, Litecoin y Monero en dólares, del 1 de enero de 2017 al 31 de diciembre de 2019	59
3	Estadísticos de resumen de la rentabilidad continua de Bitcoin, Ethereum, Ethereum Classic, Ripple, Litecoin y Monero, del 1 de enero de 2017 al 31 de diciembre de 2019	60
4	Distribuciones ajustadas y resultados de los criterios de información para el ajuste de rentabilidad continua de Bitcoin del 1 de enero de 2017 al 31 de diciembre de 2019	66
5	Distribuciones ajustadas y resultados de los criterios de información para el ajuste de rentabilidad continua de Ripple del 1 de enero de 2017 al 31 de diciembre de 2019	66
6	Distribuciones ajustadas y resultados de los criterios de información para el ajuste de rentabilidad continua de Ethereum del 1 de enero de 2017 al 31 de diciembre de 2019	67

7	Distribuciones ajustadas y resultados de los criterios de información para el ajuste de rentabilidad continua de Ethereum Classic del 1 de enero de 2017 al 31 de diciembre de 2019	67
8	Distribuciones ajustadas y resultados de los criterios de información para el ajuste de rentabilidad continua de Litecoin del 1 de enero de 2017 al 31 de diciembre de 2019	67
9	Distribuciones ajustadas y resultados de los criterios de información para el ajuste de rentabilidad continua de Monero del 1 de enero de 2017 al 31 de diciembre de 2019	68
10	Test de Kolmogorv Smirnov a dos colas para la distribución que mejor ajusta la rentabilidad continua de las 6 criptomonedas del 1 de enero de 2017 al 31 de diciembre de 2019	71
11	Precios de cierre de Bitcoin para el periodo comprendido entre el 2 de enero de 2017 y el 8 de enero de 2017.	80
12	Precios de cierre de la tabla 11 en formato supervisado.	80
13	Precios supervisados de la tabla 12 a los que se les ha aplicado la normalización de la fórmula 2.	81
14	Test ADF para las 6 criptomonedas y número de diferenciaciones aplicadas a cada una de ellas	85
15	Raíz del error cuadrático medio (calculado en el validation set) y AIC (calculado en el reduced training) de los mejores modelos de cada rango según el criterio de AIC. Para cada rango y criptomoneda se han encontrado aquellas p y q que minimizan el criterio de AIC. Una vez encontrado el mejor modelo según AIC se evalúa su ECM en el validation set para ver si de los tres modelos con menor AIC encontrados para cada criptomoneda, aquel con menor AIC es el que tiene menor ECM.	87
16	Tabla en la que se han evaluado la raíz del error cuadrático medio en la predicción de los precios del validation set con diferentes combinaciones de número de time steps y tasas de abandono.	105
17	Combinación de m y p utilizadas finalmente	105
18	ECM evaluado en el test set	106
19	Porcentaje de aciertos evaluando en el test set	111

22 Código

Todo el código mostrado a continuación es propio y ha sido programado en Jupyter Notebook mediante el lenguaje Python. Todo el trabajo se ha realizado mediante funciones para así automatizar todos los procesos. Por ejemplo, toda la parte de análisis estadístico de las criptomonedas podría ser realizada automáticamente con otras criptomonedas. ##Importaciones

```
import sys
assert sys.version_info >= (3, 5)
import sklearn
assert sklearn.__version__ >= "0.20"
import numpy as np
import os
import pandas as pd
%matplotlib inline
import matplotlib as mpl
mpl.rc('axes', labelsz=14)
mpl.rc('xtick', labelsz=12)
mpl.rc('ytick', labelsz=12)
```

```

from scipy.stats import skew, kurtosis
import scipy.stats as st
import pylab
import sys
assert sys.version_info >= (3, 5)
from matplotlib.pyplot import figure
from sklearn.model_selection import RandomizedSearchCV
from tensorflow.keras.layers import LayerNormalization
import sklearn
assert sklearn.__version__ >= "0.20"
from sklearn.preprocessing import MinMaxScaler
import itertools
import numpy as np
import os
from sklearn.metrics import mean_squared_error
from sklearn.metrics import mean_absolute_error
%matplotlib inline
import matplotlib as mpl
import matplotlib.pyplot as plt
import statsmodels.api as sm
import pylab
import gc
import datetime
from matplotlib import pyplot
import keras
from keras.models import Sequential
from keras.layers import Activation, Dense
from keras.layers import LSTM
from keras.layers import Dropout
from datetime import datetime
import time
import html5lib
import seaborn as sns
sns.set_style('whitegrid')
%matplotlib inline
import statsmodels.api as sm
from statsmodels.tsa.arima_model import ARIMA
from statsmodels.tsa.stattools import adfuller
import matplotlib.pyplot as plt
import seaborn as sns
sns.set_style('whitegrid')
%matplotlib inline
from keras.callbacks import EarlyStopping, ModelCheckpoint
from math import sqrt

```

22.1 Análisis estadístico de las criptomonedas

```

def return_close_prices (cryptos):
    """
    cryptos: list with cryptocurrencies
    Importa todas las tablas y acota el periodo a analizar al que va entre el 01/01/2017 al 30/12/2019
    """
    CloseDataFrame_price= pd.DataFrame()
    CloseDataFrame_logreturns= pd.DataFrame()
    CloseDataFrame_returns= pd.DataFrame()
    for name in cryptos:
        cryptocurrency_price = read_crypto("{}_csv".format(name)).sort_index()
        logcryptocurrency = cryptocurrency_price.copy()
        cryptocurrency_returns=cryptocurrency_price.copy()
        cryptocurrency_price[name] = cryptocurrency_price["Close"].loc[ "2016-12-31":"2019-12-31" ]
        cryptocurrency_returns[name] = cryptocurrency_price["Close"].loc[ "2016-12-31":"2019-12-31" ].pct_change()
        logcryptocurrency[name] = np.log(1+ cryptocurrency_returns[name]).loc["2016-12-31":"2019-12-31" ]
        CloseDataFrame_price= pd.concat([CloseDataFrame_price, cryptocurrency_price[name]], axis=1)
        CloseDataFrame_logreturns = pd.concat([CloseDataFrame_logreturns, logcryptocurrency[name]], axis=1)

    return CloseDataFrame_price.dropna(), CloseDataFrame_logreturns.dropna()

#####

def return_stats_of_close_prices():
    """
    Devuelve la tabla de estadísticos
    """
    format_prices = lambda x : "%.2f" % x
    format_log = lambda x : "%.4f" % x
    returns_closefinal, logreturns_closefinal= return_close_prices(cryptos)
    logreturns_descriptive = logreturns_closefinal.describe().iloc[1:]
    returns_descriptive = returns_closefinal.describe().iloc[1:]
    returns_stats= pd.DataFrame(np.array([skew(returns_closefinal),kurtosis(returns_closefinal),
        np.array(returns_descriptive.loc["max"]-returns_descriptive.loc["min"]),
        np.square(returns_descriptive.loc["std"]),
        np.array(returns_descriptive.loc["std"]/returns_descriptive.loc["mean"])]),
        index = ["Skew", "Kurtosis", "Rango","Variance", "CV"],
        columns = returns_closefinal.columns.values)
    logreturns_stats= pd.DataFrame(np.array([skew(logreturns_closefinal),kurtosis(logreturns_closefinal),
        np.array(logreturns_descriptive.loc["max"]-logreturns_descriptive.loc["min"]),
        np.square(logreturns_descriptive.loc["std"]),
        np.array(logreturns_descriptive.loc["std"]/logreturns_descriptive.loc["mean"])]),
        index = ["Skew", "Kurtosis", "Range","Variance", "CV"],

```

```

        columns = logreturns_closefinal.columns.values)
final_return = returns_descriptive.append(returns_stats)
final_logreturn = logreturns_descriptive.append(logreturns_stats)
names = ["25%", "50%", "75%", "mean", "std", "Variance", "min", "max", "Skew", "Kurtosis"]
new_names = ["Cuartil 1", "Mediana", "Cuartil 3", "Media", "Desv. Estándar", "Varianza", "Mínimo", "Máximo", "Asimetría", "Curtosis"]
as_list = final_return.index.tolist()
as_list2 = final_logreturn.index.tolist()
for i in range(len(names)):
    idx = as_list.index(names[i])
    as_list[idx] = new_names[i]
    final_return.index = as_list
    final_logreturn.index = as_list

r_final = final_return.applymap(format_prices)
lr_final = final_logreturn.applymap(format_log)
r_final.columns.names = ["Estadísticos"]
lr_final.columns.names = ["Estadísticos"]
r_final.to_csv("summary_prices.csv", encoding="utf-8-sig")
lr_final.to_csv("summary_logreturns.csv", encoding="utf-8-sig")
return r_final, lr_final

#####
def plot_close_prices(plot_returns):
    """
    plot_returns: if True then plot returns, if false then plot logreturns.
    cryptos: list with cryptocurrencies to plot.
    """
    returns_closefinal, logreturns_closefinal = return_close_prices(cryptos)
    if plot_returns:
        plot_this = returns_closefinal
    else:
        plot_this = logreturns_closefinal

    if (plot_this.shape[1] % 2 == 0):
        ncols = 2
        nrows = plot_this.shape[1] / 2
    else:
        ncols = 2
        nrows = np.ceil(plot_this.shape[1] / 2)
    plt.figure(figsize=(10, 15))
    for i in range(plot_this.shape[1]):
        plt.subplot(nrows, ncols, i + 1)
        # plot_this.iloc[:, i].plot.hist(bins=20)
        plt.hist(np.array(plot_this.iloc[:, i]), histtype="step", color="black", bins=50)
        if plot_returns:
            plt.xlabel("$Rendimientos$", fontsize=14)
        else:
            plt.xlabel("$Rendimientos$ " + "$Logarítmicos$", fontsize=14)
        plt.ylabel("Frequency", fontsize=14)
        plt.title(cryptos[i], fontsize=20)
        plt.tight_layout()
    if plot_returns:
        save_fig("ReturnsExchange")
    else:
        save_fig("LogReturnsExchange")
    plt.show()

#####
def find_best_distr_from_data(distributions):
    """
    data: dataframe with close prices of each cryptocurrency
    distributions: list with names of scipy models to test
    Devuelve la distribución que mejor ajusta los datos de la rentabilidad continua,
    el número de parámetros que ésta tiene, el número de individuos utilizados,
    devuelve la log verosimilitud y los parámetros de la distribución que mejor ajusta.
    """
    returns_closefinal, logreturns_closefinal = return_close_prices(cryptos)
    data_distributions = distributions
    best_distributions = []
    k = []
    n = []
    best_distr = []
    LLH = []
    params = []
    for i in range(logreturns_closefinal.shape[1]):
        best_AIC = 1000000000
        data = logreturns_closefinal.iloc[:, i]
        for j in range(len(data_distributions)):
            dist = getattr(st, distributions[j]).name
            params_distr = dist.fit(data)
            LLH_distr = np.sum(distributions[j].logpdf(data, *params_distr))
            k_distr = len(params_distr)
            n_distr = data.shape[0]
            AIC = 2 * k_distr + (-2 * (LLH_distr))
            if AIC < best_AIC and best_AIC == 1000000000:
                best_AIC = AIC
                best_distr.append(distributions[j].name)
                k.append(k_distr)
                n.append(n_distr)
                LLH.append(LLH_distr)
                params.append(params_distr)
            elif AIC < best_AIC:

```



```

        best_AIC=AIC
        best_distr[i] = distributions[j].name
        k[i]= k_distr
        n[i]=n_distr
        LLH[i]= LLH_distr
        params[i]=params_distr

    return best_distr, k, n, LLH,params

#####

def return_information_criteria():
    """
    Devuelve tabla con criterios de información.
    """
    distr = distributions
    returns_closefinal, logreturns_closefinal= return_close_prices(cryptos)
    for i in range(logreturns_closefinal.shape[1]):
        information_criteria = np.empty((len(distributions), 6))
        data=data=logreturns_closefinal.iloc[:,i]
        for j in range(len(distr)):
            dist = getattr(st,distributions[j].name)
            params = dist.fit(data)
            LLH_distr= dist.logpdf(data,'params').sum()
            LLH_negative= -LLH_distr
            k_distr = len(params)
            n_distr= data.shape[0]
            AIC= 2*k_distr - 2*(LLH_distr)
            BIC= k_distr*np.log(n_distr) - 2*LLH_distr
            CAIC = -2*LLH_distr + k_distr *(np.log(n_distr)+1)
            AICc= AIC + ((2*k_distr*(k_distr+1))/(n_distr-k_distr-1))
            HQC = -2*LLH_distr + 2*k_distr*(np.log(np.log(n_distr)))
            inf = np.array([LLH_negative,AIC,BIC,CAIC,AICc,HQC])
            information_criteria[j,:]= inf
        if not os.path.exists("inform_{}.csv".format(logreturns_closefinal.columns.values[i])):
            df = pd.DataFrame(information_criteria, index = [i.name for i in distributions],columns = criterios)
            df.columns.names=["Distribution"]
            df.to_csv("inform_{}.csv".format(logreturns_closefinal.columns.values[i]))
        else:
            print("Already have {}".format(logreturns_closefinal.columns.values[i]))

#####
def fit_of_best_distributions2(distributions):
    """
    Devuelve el test de kolmogorov para la distribución que mejor ajusta los datos de la rentabilidad continua.
    """
    best_distr, k, n, LLH,params = find_best_distr_from_data(distributions)
    returns_closefinal,logreturns_closefinal= return_close_prices(cryptos)
    ncol = logreturns_closefinal.shape[1]
    fits= np.empty((6,))
    for i in range(ncol):
        ks= st.kstest(logreturns_closefinal.iloc[:,i], best_distr[i],args=(params[i]))
        ks_p_value=ks[1]
        fits[i] = ks_p_value
    df = pd.DataFrame(fits, index=logreturns_closefinal.columns.values,columns =["KS P-value"])
    #if shape is (6,) then it must have one column
    df.columns.names=["Cryptocurrency"]
    df.to_csv("fit_cryptos")
    return df

#####
def list_parameters(distribution):
    """
    Devuelve los parámetros que una distribución tiene dado el nombre de la distribución
    """
    if isinstance(distribution, str):
        distribution = getattr(st, distribution)
    if distribution.shapes:
        parameters = [name.strip() for name in distribution.shapes.split(',')]
    else:
        parameters = []
    if distribution.name in st._discrete_distns._distn_names:
        parameters += ['loc']
    elif distribution.name in st._continuous_distns._distn_names:
        parameters += ['loc', 'scale']
    else:
        sys.exit("Distribution name not found in discrete or continuous lists.")
    return parameters

def return_dictionary_with_parameters():
    """
    Devuelve dos vectores, uno con los nombres de los parámetros (a,b, loc y scale) y el
    valor asociado a cada uno para cada criptomoneda.
    """
    best_distr, k, n, LLH,params = find_best_distr_from_data(distributions)
    par_names=[]
    par_numbers=[]
    for i in range(len(best_distr)):
        names =list_parameters(best_distr[i])
        par_names.append(names)
        par_numbers.append(params[i])
    return par_names, par_numbers

```

```
#####
def graph_VaRs (distributions,plot_returns):
    """
    Realiza la figura 25 con plot_returns=False.
    """
    best_distr, k, n, LLH,params = find_best_distr_from_data(distributions)
    returns_closefinal,logreturns_closefinal= return_close_prices(cryptos)
    vars_crypto= np.empty((3,logreturns_closefinal.shape[1]))
    distribution_names = [names.name for names in distributions]
    plt.figure(figsize=(10,15))
    ab=np.arange(0.0,1,0.01)
    if plot_returns:
        plot_this = returns_closefinal
    else:
        plot_this = logreturns_closefinal

    if(plot_this.shape[1] % 2 ==0):
        ncols= 2
        nrows = plot_this.shape[1]/2
    else:
        ncols=2
        nrows = np.ceil(plot_this.shape[1]/2)
    plt.figure(figsize=(10,15))
    for i in range(len(best_distr)):
        plt.subplot(nrows,ncols,i+1)
        VaRs= distributions[distribution_names.index(best_distr[i])].ppf(ab,*params[i])
        plt.plot(ab,VaRs ,color="black", linewidth=2,ls="--")
        plt.xlabel("$q$", fontsize=14)
        plt.ylabel("$VaR(q)$", fontsize=14)
        plt.title(cryptos[i], fontsize=20)
        plt.tight_layout()
    save_fig("VaR")
    plt.show()

#####

def graph_CVaR (distributions,plot_returns):
    """
    Realiza la figura 26
    """
    best_distr, k, n, LLH,params = find_best_distr_from_data(distributions)
    returns_closefinal,logreturns_closefinal= return_close_prices(cryptos)
    vars_crypto= np.empty((3,logreturns_closefinal.shape[1]))
    distribution_names = [names.name for names in distributions]
    plt.figure(figsize=(10,15))
    ab=np.arange(0.01,1,0.01)
    if plot_returns:
        plot_this = returns_closefinal
    else:
        plot_this = logreturns_closefinal

    if(plot_this.shape[1] % 2 ==0):
        ncols= 2
        nrows = plot_this.shape[1]/2
    else:
        ncols=2
        nrows = np.ceil(plot_this.shape[1]/2)
    plt.figure(figsize=(10,15))
    for i in range(len(best_distr)):
        plt.subplot(nrows,ncols,i+1)
        bootstrap_p= distributions[distribution_names.index(best_distr[i])].rvs("params[i]", size=1000000)
        VaRs= distributions[distribution_names.index(best_distr[i])].ppf(ab,*params[i])
        data_cvar= np.empty((len(VaRs),1))
        for j in range(len(VaRs)):
            data_cvar[j] =CVaRs= sum(bootstrap_p[bootstrap_p<=VaRs[j]])/len(bootstrap_p[bootstrap_p<=VaRs[j]])
        plt.plot(ab,data_cvar ,color="black", linewidth=2,ls="--")
        plt.xlabel("$VaR(q)$", fontsize=14)
        plt.ylabel("$CVaR(q)$", fontsize=14)
        plt.title(cryptos[i], fontsize=20)
        plt.tight_layout()
    save_fig("CVaR")
    plt.show()

#####

def make_pdf(dist, params, size=10000):
    """Devuelve una serie con los datos necesarios para hacer el gráfico de la distribución de probabilidad
    ajustada a los datos de la rentabilidad continua """
    arg = params[:-2]
    loc = params[-2]
    scale = params[-1]

    start = dist.ppf(0.01, *arg, loc=loc, scale=scale) if arg else dist.ppf(0.01, loc=loc, scale=scale)
    end = dist.ppf(0.99, *arg, loc=loc, scale=scale) if arg else dist.ppf(0.99, loc=loc, scale=scale)

    x = np.linspace(start, end, size)
    y = dist.pdf(x, loc=loc, scale=scale, *arg)
    pdf = pd.Series(y, x)

    return pdf

def plot_density_function(distributions, plot_returns=False):
```

"""Realiza la figura 22"""

```
best_distr, k, n, LLH, params = find_best_distr_from_data(distributions)
returns_closefinal, logreturns_closefinal = return_close_prices(cryptos)
distribution_names = [names.name for names in distributions]
if plot_returns:
    plot_this = returns_closefinal
else:
    plot_this = logreturns_closefinal

if (plot_this.shape[1] % 2 == 0):
    ncols = 2
    nrows = plot_this.shape[1] / 2
else:
    ncols = 2
    nrows = np.ceil(plot_this.shape[1] / 2)
plt.figure(figsize=(12, 15))
for i in range(logreturns_closefinal.shape[1]):
    plt.subplot(nrows, ncols, i + 1)
    data = logreturns_closefinal.iloc[:, i]
    pdf = make_pdf(distributions[distribution_names.index(best_distr[i])], params[i])
    best_distr = getattr(st, best_distr[i])
    best_fit_params = params[i]
    best_fit_name = distributions[distribution_names.index(best_distr[i])].name
    ax = pdf.plot(lw=2, label="PDF", legend=True, color="black")
    data.plot(kind="hist", histtype="step", color="black", bins=75, density=True, alpha=0.5, label="Data", legend=True, ax=ax)
    param_names = (best_dist.shapes + ', loc, scale').split(',') if best_dist.shapes else ['loc', 'scale']
    param_str = ', '.join(['{0:0.2f}'.format(k, v) for k, v in zip(param_names, best_fit_params)])
    dist_str = '{0()}'.format(best_fit_name, param_str)
    ax.set_title(logreturns_closefinal.columns.values[i] + "n" + dist_str)
    ax.set_xlabel("Rendimientos" + " logarítmicos")
    ax.set_ylabel("Frecuencia")
save_fig("Prob_fit_distr")
plt.show()
```

22.2 Modelo ARIMA y redes LSTM

MODELO ARIMA

```
def compute_ADF_test(close_prices):
    """
    Realiza el test ADF en cada una de las criptomonedas
    """
    close_prices2 = close_prices.copy()
    ncol = close_prices2.shape[1]
    n_dif = 5
    fits = np.empty((ncol, 5))
    diferenciaciones = np.empty((ncol))
    for i in range(ncol):
        dif = 0
        close_prices2[close_prices2.columns[i] + "2"] = close_prices2.iloc[:, i]
        ADF = adfuller(close_prices2[close_prices2.columns[i] + "2"], regression="ct", autolag="t-stat")
        if ADF[0] <= ADF[4][0.05]:
            fits[i, 0] = ADF[0]
            fits[i, 1] = ADF[4][0.01]
            fits[i, 2] = ADF[4][0.05]
            fits[i, 3] = ADF[4][0.10]
            fits[i, 4] = dif
            diferenciaciones[i] = dif
        else:
            for j in range(n_dif):
                dif += 1
                close_prices2[close_prices2.columns[i] + "2"] = close_prices2[close_prices2.columns[i] + "2"].diff()
                close_prices2 = close_prices2.dropna()
                ADF = adfuller(close_prices2[close_prices2.columns[i] + "2"], regression="ct", autolag="t-stat")
                if ADF[0] <= ADF[4][0.05]:
                    fits[i, 0] = ADF[0]
                    fits[i, 1] = ADF[4][0.01]
                    fits[i, 2] = ADF[4][0.05]
                    fits[i, 3] = ADF[4][0.10]
                    fits[i, 4] = dif
                    diferenciaciones[i] = dif
                    break

df = pd.DataFrame(fits, index=close_prices2.columns.values[ncol:], columns=
    ["Estadístico test", "Valor crítico 1 %", "Valor crítico 5 %", "Valor crítico 10 %", "n° de dif."])
df["n° de dif."] = df["n° de dif.".astype('int')]
df.columns.names = ["Criptomoneda"]
df.to_csv("ADF_test.csv", encoding="utf-8-sig")
return df, diferenciaciones

def best_params_ARIMA(close_prices, diferenciaciones, max_pq):
    """
    Si a esta función le das los precios de cierre, el número de diferenciaciones a aplicar a
    cada criptomoneda y la máxima p y q que pueden ser utilizadas los parámetros del mejor
    modelo ARIMA atendiendo al criterio de AIC.
    """
    params = []
    params_final = []
    AIC = []
```

```

train,val,test= define_train_and_test(close_prices, 0.8,0.6)
for i in range(close_prices.shape[1]):
    training_crypto = train.iloc[:,i]
    p=q=range(0,max_pq)
    d= range(int(differenciaciones[i]), int(differenciaciones[i]+1))
    pdq = [(x[0], x[1], x[2]) for x in list(itertools.product(p, d, q))]
    warnings.filterwarnings("ignore")
    AIC_list = pd.DataFrame({}, columns=['param','AIC'])
    for param in pdq:
        try:
            if max_pq==10:
                print(param)
            mod = sm.tsa.statespace.SARIMAX(training_crypto,order=param,seasonal_order=(0, 0, 0, 0))
            results = mod.fit()
            temp = pd.DataFrame([ param , results.aic ], columns=['param','AIC'])
            AIC_list = AIC_list.append( temp, ignore_index=True)
            del temp
        except:
            continue

    m = np.amin(AIC_list['AIC'].values) # Find minimum value in AIC
    l = AIC_list['AIC'].tolist().index(m) # Find index number for lowest AIC
    Min_AIC_list = AIC_list.iloc[l,:]
    params.append(Min_AIC_list)
    for k in range(len(params)):
        params_final.append(params[k][0])
    for l in range(len(params)):
        AIC.append(params[l][1])
    return params_final, AIC

def flatten(column):
    """
    Elimina los NA
    """
    return column.dropna().reset_index(drop=True)

def sqrt_eqm(valores_reales,pred):
    """
    Calcula el error cuadrático medio
    """
    result= sqrt(((valores_reales-pred) ** 2).mean())
    return result

def dataframe_of_params(best_params,AIC, close_prices):
    """
    Transforma los datos en un dataframe para crear la tabla 15
    """
    format_x= lambda x : int(x)
    values = np.empty((close_prices.shape[1], 5))
    for j in range(close_prices.shape[1]):
        for i in range(4):
            if i!=3:
                values[j,i] = best_params[j][i]
            else:
                values[j,i] = AIC[j]
    data = pd.DataFrame(data=values, index=close_prices.columns.values, columns=["p", "d", "q", "AIC", "Raiz_ECM"]).applymap(format_x)
    return data

def accuracy_arima(n_init_training,dataset, predicted, expected):
    """
    Calcula el accuracy tanto para los modelos ARIMA como para la red LSTM a partir de sus predicciones. La accuracy es el porcentaje de veces que ha predicho correctamente la evolución de los precios.
    """
    vector = []
    predictions= []
    expected_updown=[]
    if isinstance(dataset, pd.DataFrame):
        dataset = np.array(dataset)
    if isinstance(predicted, pd.DataFrame):
        predicted = np.array(predicted)
    if isinstance(expected, pd.DataFrame):
        expected = np.array(expected)
    last_value_training = dataset[n_init_training-1]
    print(len(predicted), len(expected))
    if (len(predicted) != len(expected)):
        print("shapes not equal")
        return
    for i in range(len(predicted)):
        if(i ==0):
            if (predicted[i]>last_value_training) and (expected[i]>last_value_training):
                vector.append(1)
                predictions.append(1)
                expected_updown.append(1)
            elif (predicted[i]<last_value_training) and (expected[i]<last_value_training):
                vector.append(1)
                predictions.append(-1)
                expected_updown.append(-1)
            elif (predicted[i]>last_value_training):
                vector.append(0)
                predictions.append(1)
                expected_updown.append(-1)

```

```

else:
    vector.append(0)
    predictions.append(-1)
    expected_updown.append(1)
    print("predicted", predicted[i], "expected", expected[i], "last", last_value_training)
else:
    if (predicted[i]>expected[i-1]) and (expected[i]>expected[i-1]):
        vector.append(1)
        predictions.append(1)
        expected_updown.append(1)
    elif (predicted[i]<expected[i-1]) and (expected[i]<expected[i-1]):
        vector.append(1)
        predictions.append(-1)
        expected_updown.append(-1)
    elif (predicted[i]>expected[i-1]):
        vector.append(0)
        predictions.append(1)
        expected_updown.append(-1)
    else:
        vector.append(0)
        predictions.append(-1)
        expected_updown.append(1)
    print("predicted", predicted[i], "expected", expected[i], "last", expected[i-1])

pred=np.array(predictions)
exp = np.array(expected_updown)
all_pe= np.c_[pred,exp]

return (np.sum(np.array(vector))/len(np.array(vector)))*100,all_pe

def predict_EQM_ARIMA_on_validation(data= close_prices):
    """
    Permite crear la tabla 15 a partir de las funciones anteriores. Para estimar en el test set
    se ha hecho algo parecido solo que cogiendo los parámetros que dan el menor ECM en el
    validation set para cada criptomoneda, ECM presente en la tabla 15.
    """
    format_EQM= lambda x : "%.4f" % x
    max_pq = np.array([4,6,10])
    sqrt_EQM = np.empty((len(max_pq),close_prices.shape[1]))
    nforecasts = 1
    tickers= ["np=3,nq=3", "np=5,nq=5", "np=9,nq=9"]

    for i in range(len(max_pq)):
        print("hola1")
        best_params, AIC = best_params_ARIMA(close_prices, differenciations,max_pq[i])
        print("hola2")
        if i==0:
            data1= dataframe_of_params(best_params,AIC, close_prices)
        elif i==1:
            data2= dataframe_of_params(best_params,AIC, close_prices)
        else:
            data3= dataframe_of_params(best_params,AIC, close_prices)

    for j in range(close_prices.shape[1]):
        print(i,j)
        # Setup forecasts
        forecasts = {}
        train, val, test= define_train_and_test(pd.DataFrame(close_prices.iloc[:,j]), 0.8, 0.6)
        mod = sm.tsa.SARIMAX(train, order=best_params[j])
        res = mod.fit()
        n_training = train.shape[0]
        n_obs= train.shape[0]+ val.shape[0]
        # Step through the rest of the sample
        for t in range(n_training, n_obs):
            # Update the results by appending the next observation
            updated_endog = close_prices.iloc[t+1,j]
            # Save the new set of forecasts
            forecasts[updated_endog.index[0]] = res.forecast(steps=nforecasts)
            res = res.extend(updated_endog)
        # Combine all forecasts into a dataframe
        forecasts = pd.concat(forecasts, axis=1)
        predicciones = forecasts.apply(flatten)
        predicciones.index = (predicciones.index + 1).rename('horizon')
        predicciones2= predicciones.T
        valores_reales = np.array(val)
        pred = np.array(predicciones2)
        if i==0:
            data1.iloc[j, 4] = sqrt_eqm(valores_reales,pred)
        elif i==1:
            data2.iloc[j, 4] = sqrt_eqm(valores_reales,pred)
        else:
            data3.iloc[j, 4] = sqrt_eqm(valores_reales,pred)
    params = pd.concat([data1,data2,data3], keys=tickers, axis=1)
    params2= params.T
    params2.to_csv("pq_params.csv",encoding="utf-8-sig")
    return params2

```

RED LSTM

```

def series_to_supervised(data, n_in=1, n_out=1, dropnan=True):
    """
    Transforma los datos para que tengan el formato de la tabla 12.
    """
    n_vars = 1 if type(data) is list else data.shape[1]
    df = pd.DataFrame(data)
    cols, names = list(), list()
    # input sequence (t-n, ... t-1)
    for i in range(n_in, 0, -1):
        cols.append(df.shift(i))
        names += [('Precio(t-%d)' % (i)) ]
    # forecast sequence (t, t+1, ... t+n)
    for i in range(0, n_out):
        cols.append(df.shift(-i))
        if i == 0:
            names += [('Precio(t)')]
        else:
            names += [('Precio%d(t+%d)' % (i+1, i)) for j in range(n_vars)]
    # put it all together
    agg = pd.concat(cols, axis=1)
    agg.columns = names
    agg = agg.iloc[n_in:]
    return agg

def normalize_sequences(data):
    """
    Transforma la tabla 12 en la tabla 13
    """
    copy = data.copy()
    for i in range(copy.shape[0]):
        first_element=copy.iloc[i,0]
        for j in range(copy.shape[1]):
            copy.iloc[i,j] = (copy.iloc[i,j]/first_element) -1
    return(copy)

def inverse_normalize(data_original, data_transform_Y,row):
    """
    Invierte la normalización de las predicciones para así tener el precio de cierre predicho
    """
    print("prediction_raw", data_transform_Y)
    new_value = (data_transform_Y+1)*data_original.iloc[row,0]
    array = np.array(new_value)
    return(new_value)

def inverse_and_normalize_and_define_train_val_test (data, steps, acc_xy):
    """
    A partir de los precios de cierre, convierte el problema en supervisado, normaliza los
    datos, define el training, validation y test set y concatena el training y el validation
    formando train_val, que es utilizado cuando se evalúa la capacidad predictiva final en el
    test set. Cuando se evalúa en el validation set se utiliza el dataframe llamado train.
    """
    if acc_xy:
        sup_data = series_to_supervised(data,steps)
        norm_sup_data = normalize_sequences(sup_data)
        train, val, test= define_train_and_test(norm_sup_data, 0.8, 0.6)
        train_val= pd.concat([train,val], axis=0)
    else:
        sup_data = series_to_supervised(data,steps)
        norm_sup_data = normalize_sequences(sup_data)
        train, val, test= define_train_and_test(norm_sup_data, 0.8, 0.6)
        train_val= pd.concat([train,val], axis=0)
    return train, val, test, train_val

def build_model_2003(eval_on_val, rate,n_neurons, train_lstm, val_lstm,test_lstm, train_val_lstm, steps,acc_xy,batch_size):
    """
    Aquí se define el modelo, y se le proporciona el número de neuronas que tendrá, la dimensionalidad, si va a evaluar
    en el validation set o en el test set y sobre la tasa de caída de neuronas de la red LSTM, definida como rate.
    """
    if eval_on_val:
        X_train, y_train = obtain_values_x_y(train_lstm,steps, acc_xy)
    else:
        X_train, y_train = obtain_values_x_y(train_val_lstm,steps,acc_xy)

    print("batch",batch_size)
    model4 = keras.models.Sequential([
        keras.layers.LSTM(n_neurons, batch_input_shape=(batch_size, X_train.shape[1], X_train.shape[2])),
        # [batch_size(it can vary, every batch the states are reset), time_steps (it can vary), dimensionality (1 variable)]
        keras.layers.Dropout(rate),
        keras.layers.Dense(1)])
    model4.compile(loss="mse", optimizer="adam")
    print(model4.summary())
    return model4

def obtain_values_x_y(data,steps,acc_xy):
    """
    Crea dos arrays, uno con los datos con los que se entrena la red LSTM y el otro con los
    valores que tiene que tratar de predecir a partir de los datos con los que se entrena. No
    hacer caso a acc_xy cuando es True ya que esta por defecto ha estado como False.
    """
    if acc_xy:
        if isinstance(data, pd.DataFrame):

```

```

    data = np.array(data)
    print(data[0:3,:])
    a= data[:, 0:-1]
    a= a.reshape(a.shape[0], steps,1)
    b=data[:,1:]>=data[:,2]
    b= np.array([int(x) for x in b])
    return a,b
else:
    if isinstance(data, pd.DataFrame):
        data = np.array(data)
        a= data[:, 0:-1]
        a= a.reshape(a.shape[0], steps,1)
        b = data[:, -1]
    return a ,b

def make_all_forecasts_dif_size(data,steps, forecast_on_val, batch_size, batch_size_predict,model,new_model, train_lstm, val_lstm, test_lstm, train_val_lstm,i):
    """
    Aqui es donde se realizan todos los ajustes de los modelos en todo el training set
    (reduced training set + validation set) y se llevan a cabo las predicciones en el test
    set. forecast_on_val está por defecto en falso ya que hay otra función prácticamente
    igual a esta que se encarga de hacer los ajustes en el reduced training set y evaluar en
    el validation set. La única diferencia respecto a la otra es que esta no tiene early stopping.
    """
    acc_xy=False
    if forecast_on_val:
        X_train, y_train = obtain_values_x_y(train_lstm,steps,acc_xy)
        train_elements = train_lstm.shape[0]
        predict_elements=val_lstm.shape[0]
        X_predict_on, y_predict_on=obtain_values_x_y(val_lstm,steps,acc_xy)
    else:
        X_train, y_train = obtain_values_x_y(train_val_lstm,steps,acc_xy)
        train_elements = train_lstm.shape[0]+val_lstm.shape[0]
        predict_elements=test_lstm.shape[0]
        X_predict_on, y_predict_on =obtain_values_x_y(test_lstm,steps,acc_xy)
    predictions = list()
    train_original, val_original, test_original = define_train_and_test(data, 0.8, 0.6)
    steps_cripto = series_to_supervised(data,steps)
    steps_cripto_train,steps_cripto_val, steps_cripto_test = define_train_and_test(steps_cripto, 0.8, 0.6)
    print("obs",batch_size,X_train.shape, y_train.shape, X_predict_on.shape, y_predict_on.shape)
    model.fit(X_train, y_train, epochs=75, batch_size=batch_size, shuffle=False)
    old_weights = model.get_weights()
    new_model.set_weights(old_weights)
    train_lstm,val_lstm,test_lstm, train_val_lstm=inverse_and_normalize_and_define_train_val_test(data, steps, acc_xy)
    if forecast_on_val:
        X_train, y_train = obtain_values_x_y(train_lstm,steps,acc_xy)
        train_elements = train_lstm.shape[0]
        predict_elements=val_lstm.shape[0]
        X_predict_on, y_predict_on =obtain_values_x_y(val_lstm,steps,acc_xy)
    else:
        X_train, y_train = obtain_values_x_y(train_val_lstm,steps,acc_xy)
        train_elements = train_lstm.shape[0]+val_lstm.shape[0]
        predict_elements=test_lstm.shape[0]
        X_predict_on, y_predict_on =obtain_values_x_y(test_lstm,steps,acc_xy)
    for i in range(predict_elements): #here we make predictions one by one and invert all the transformations
        # make one-step forecast
        X_forecast = X_predict_on[i,:]
        row= train_elements + i #train
        yhat = forecast_lstm(new_model, steps,X_forecast,batch_size_predict) #I give the last values
        yhat= inverse_normalize(steps_cripto, yhat,row) #it uses the first column
        predictions.append(yhat)
    if forecast_on_val:
        expected= steps_cripto_val.iloc[:,1].values
        ab= np.c_[predictions, expected]
        dataf = pd.DataFrame(ab, columns=["Predicted", "Expected"], index= val_original.index)
        acc, pred_up_down = accuracy_arma(train_elements,data, predictions, expected)
        pred_updown = pd.DataFrame(pred_up_down, columns=["Predicted", "Expected"], index= test_original.index)
    else:
        expected= steps_cripto_test.iloc[:,1].values
        ab= np.c_[predictions, expected]
        dataf = pd.DataFrame(ab, columns=["Predicted", "Expected"], index= test_original.index)
        print("train_elements", train_elements)
        acc, pred_up_down = accuracy_arma(train_elements+steps,data, predictions, expected)
        pred_updown = pd.DataFrame(pred_up_down, columns=["Predicted", "Expected"], index= test_original.index)
    # report performance
    rmse= sqrt_eqm(expected, predictions)
    print("Test RMSE: %.3f" % rmse)
    return dataf, rmse,acc, pred_updown

def dataframe_of_lstm_dif_size(steps, forecast_on_val, batch_size,batch_size_predict,neurons, drop, data,i, acc_xy):
    """
    Aqui es donde se definen los modelos. Se eliminan algunos individuos ya que el total de
    individuos con el que se entrena tiene que ser múltiple del batch size que en este caso
    es igual a 20.
    """
    values = np.empty((1,))
    train_lstm,val_lstm,test_lstm, train_val_lstm=inverse_and_normalize_and_define_train_val_test(pd.DataFrame(data.iloc[:,i]), steps, acc_xy)
    if batch_size==10:
        if steps==10:
            train_val_lstm= train_val_lstm.iloc[6:]
        if steps==5 or steps==15:
            train_val_lstm= train_val_lstm.iloc[1:]

```

```

        if batch_size==20:
            if steps==10:
                train_val_lstm= train_val_lstm.iloc[6:]
            if steps==5:
                train_val_lstm= train_val_lstm.iloc[11:]
            if steps==15:
                train_val_lstm= train_val_lstm.iloc[1:]
        print("nw", train_val_lstm.shape)
        print("drop", drop, "neurons", neurons, "steps",steps)
        model= build_model_2003(forecast_on_val,drop,neurons, train_lstm, val_lstm,test_lstm, train_val_lstm,steps, acc_xy, batch_size)
        new_model = build_model_2003(forecast_on_val,drop,neurons, train_lstm, val_lstm,test_lstm, train_val_lstm,steps, acc_xy, batch_size_predict)
        together, rmse, acc, pred_up_down=make_all_forecasts_dif_size(pd.DataFrame(data.iloc[:,i]),steps, forecast_on_val, batch_size,batch_size_predict,model,new_model,train_lstm,val_lstm,test_lstm)
        values=rmse
        return pd.Series(values), together,acc, pred_up_down

def ecm_lstm_table_test(forecast_on_val,data,batch_size, batch_size_predict,acc_xy):
    """
    Aqui se itera sobre cada criptomoneda. Esta función, para cada criptomoneda,
    llama a la función de arriba que a la vez llama la función que tiene encima.
    """

    format_EQM= lambda x : "%.4f" % x
    nforecasts = 1
    best_params_lstm= select_best_params_lstm(all_sqe_lstm)
    dataf2=pd.DataFrame()
    vector=[]
    data_accuracy = np.empty((close_prices.shape[1], 1))
    dataf3=pd.DataFrame()
    for i in range(data.shape[1]):
        drop=best_params_lstm.iloc[1,i]
        steps=best_params_lstm.iloc[0,i]
        if steps==10.00:
            neurons=10
        if steps==5.00:
            neurons=5
        if steps==15.00:
            neurons=15
        values, dataf,acc, pred_up_down= dataframe_of_lstm_dif_size(int(steps),forecast_on_val, batch_size,batch_size_predict, neurons,drop, data,i,acc_xy)
        vector.append(values)
        dataf3=pd.concat([dataf3,pred_up_down], axis=1)
        dataf2= pd.concat([dataf2,dataf], axis=1)
        data_accuracy[i,0] = acc
    sqrt_ecm=pd.concat(vector, axis=1)
    sqrt_ecm.columns=data.columns.values
    sqrt_ecm.index=["LSTM"]
    complete_dataframe = sqrt_ecm
    complete2= complete_dataframe.applymap(format_EQM).T
    accuracy = pd.DataFrame(data_accuracy, index= close_prices.columns.values, columns=["LSTM"])
    accuracy_export= save_table (accuracy,"accuracy7_lstm_test.csv", path=tables_path)
    sqrt_lstm= save_table (complete2,"sqrt4_ecm_lstm_test.csv", path=tables_path)
    predictions_lstm= save_table (dataf2,"predictions4_lstm_test.csv", path=tables_path)
    predictions_updown= save_table (dataf3,"updown4_lstm_test.csv", path=tables_path)
    return complete2, dataf2, accuracy, dataf3

#####

```

23 Bibliografía

23.1 Fuente de los datos

BNC Digital Currency Indexed EOD [consulta 16 de febrero de 2019]. Disponible en: <https://www.quandl.com/data/BNC2-BNC-Digital-Currency-Indexed-EOD>

The Global Findex Database 2017 [Consulta 10 de abril de 2020]. Disponible en: <https://www.betterthancash.org/tools-research/resources/the-global-findex-database-2017>

The World Bank, Migration and Remittances Data [Consulta 10 de abril de 2020]. Disponible en: <https://www.worldbank.org/en/topic/migrationremittancesdiasporaissues/brief/migration-remittances-data>

The World Bank, Remittance Prices Worldwide [Consulta 10 de abril de 2020]. Disponible en <http://remittanceprices.worldbank.org>

23.2 Fuentes Multimedia

Wikipedia [consulta 16 de febrero de 2020]. Disponible en: https://es.wikipedia.org/wiki/Prueba_de_Kolmogorov-Smirnov

Wikipedia [consulta 7 de abril de 2020]. Disponible en: https://es.wikipedia.org/wiki/Acuerdos_de_Bretton_Woods#Contenido

Wikipedia [consulta 16 de febrero de 2020]. Disponible en: https://es.wikipedia.org/wiki/Valor_en_riesgo

Wikipedia [Consulta 4 de abril de 2020]. Disponible en: <https://www.cnet.com/news/coronavirus-unemployment-payments-applying-and-more/>

Wikipedia [Consulta 10 de abril de 2020]. Disponible en https://es.wikipedia.org/wiki/Ley_de_Gresham

Wikipedia [Consulta 10 de abril de 2020]. Disponible en <https://en.wikipedia.org/wiki/Bimetallism>

Wikipedia [Consulta 11 de abril de 2020]. Disponible en https://es.wikipedia.org/wiki/Hiperinflaci%C3%B3n#cite_note-inflacionnov-34

Wikipedia [Consulta 11 de abril de 2020]. Disponible en <https://es.wikipedia.org/wiki/Desapalancamiento>

Wikipedia [Consulta 11 de abril de 2020]. Disponible en https://es.wikipedia.org/wiki/Teor%C3%ADa_cuantitativa_del_dinero

Wikipedia [Consulta 12 de abril de 2020]. Disponible en https://es.wikipedia.org/wiki/Prestamista_de_%C3%BAltima_instancia

Wikipedia [Consulta 14 de abril de 2020]. Disponible en <https://es.wikipedia.org/wiki/M-Pesa>

Blockchain Privada [Consulta 4 de abril de 2020]. Disponible en: https://academy.bit2me.com/cuantos-tipos-de-blockchain-hay/#Blockchain_privada_o_permisionada

Caia de Conversion [Consulta 7 de abril de 2020]. Disponible en: <https://www.dinero.com/archivo/articulo/caia-conversion/18566>

Diferencia entre criptomonedas, tokens, monedas virtuales y digitales [Consulta 27 de abril de 2020]. Disponible en <https://www.cripto247.com/educacion/cual-es-la-diferencia-entre-criptomonedas-token>

Dinero electrónico [Consulta 25 de abril de 2020]. Disponible en <https://economipedia.com/definiciones/dinero-electronico-2.html>

Economics and the Libra Reserve [Consulta 25 de abril de 2020]. Disponible en: <https://libra.org/en-US/economics-and-the-reserve/?noredirect=en-US#the-libra-reserve-and-protections>

Ethereum-classic [consulta 16 de febrero de 2019]. Disponible en: <https://www.brokeronline.es/ethereum-classic/>

Litecoin [consulta 16 de febrero de 2020]. Disponible en: <https://litecoin.org/es/>

Libra [consulta 4 de abril de 2020]. Disponible en <https://libra.org/es-LA/white-paper/#introduction>

Libra, la criptomoneda de Facebook [Consulta 26 de abril de 2020]. Disponible en <https://www.caixabankresearch.com/libra-la-criptomoneda-de-facebook>

Facebook's Libra [Consulta 13 de abril de 2020]. Disponible en <https://www.cnbc.com/2019/12/27/swiss-president-says-facebooks-cryptocurrency-project-libra-failed.html>

Facebook Simba [Consulta 12 de abril de 2020]. Disponible en <https://www.businessinsider.es/facebook-desplegara-cable-submarino-africa-nombre-simba-405165>

Gradiente Descendiente para aprendizaje automático [Consulta 24 de abril de 2020]. Disponible en <https://iartificial.net/gradiente-descendiente-para-aprendizaje-automatico/>

Qué es un 'token [Consulta 27 de abril de 2020]. Disponible en <https://www.bbva.com/es/que-es-un-token-y-para-que-sirve/>

Remittance prices [Consulta 19 de abril de 2020]. Disponible en <https://remittanceprices.worldbank.org/en/methodology>

Ripple [consulta 16 de febrero de 2020]. Disponible en: <https://btcdirect.eu/es-es/acerca-de-ripple>

Sweden [consulta 9 de abril de 2020]. Disponible en: <https://www.theguardian.com/money/2019/mar/09/sweden-how-cash-became-more-trouble-than-its-worth>

Masa Monetaria [Consulta 11 de abril de 2020]. Disponible en : <https://blogs.imf-formacion.com/blog/mba/masa-monetaria-dinero/>

Monero [consulta 16 de febrero de 2020]. Disponible en: <https://www.getmonero.org/>

The four different types of money [consulta 8 de abril de 2020]. Disponible en <https://quickonomics.com/different-types-of-money/>.

Token Money [consulta 8 de abril de 2020]. Disponible en <https://financial-dictionary.thefreedictionary.com/Token+Money>

Understanding RNN and LSTM [consulta 24 de abril de 2020]. Disponible en <https://towardsdatascience.com/understanding-rnn-and-lstm-f7cdf6dfc14e>

Vodacom and Safaricom [Consulta 14 de abril de 2020]. Disponible en <https://www.vodafone.com/news-and-media/vodafone-group-releases/news/vodacom-and-safaricom-joint-venture-to-accelerate-r>

23.3 Fuentes de artículos académicos y libros

Athey, Susan, and Guido W Imbens. 2019. "Machine Learning Methods That Economists Should Know About." *Annual Review of Economics* 11. Annual Reviews: 685–725.

Azari, Amin. 2019. "Bitcoin Price Prediction: An Arima Approach." *arXiv Preprint arXiv:1904.05315*.

Azzalini, Adelchi. 1985. "A Class of Distributions Which Includes the Normal Ones."

Scandinavian Journal of Statistics. JSTOR, 171–78.

Barndorff-Nielsen, Ole E. 1997. “Normal Inverse Gaussian Distributions and Stochastic Volatility Modelling.” *Scandinavian Journal of Statistics* 24 (1). Wiley Online Library: 1–13.

Chan, Stephen, Jeffrey Chu, Saralees Nadarajah, and Joerg Osterrieder. 2017. “A Statistical Analysis of Cryptocurrencies.” *Journal of Risk and Financial Management* 10 (2). Multidisciplinary Digital Publishing Institute: 12.

Chu, Jeffrey, Saralees Nadarajah, and Stephen Chan. 2015. “Statistical Analysis of the Exchange Rate of Bitcoin.” *PloS One* 10 (7). Public Library of Science.

Danielsson, Jon. 2019. “Cryptocurrencies: Policy, Economics and Fairness.” *Systemic Risk Centre Discussion Paper* 86: 2018.

De Laplace, Pierre S. 1774. “Mémoire Sur La Probabilité Des Causes Par Les événements.” *Mém. de Math. et Phys. Présentés à L’Acad. Roy. Des Sci* 6: 621–56.

Demirguc-Kunt, Asli, Leora Klapper, Dorothe Singer, Saniya Ansar, and Jake Hess. 2018. *The Global Findex Database 2017: Measuring Financial Inclusion and the Fintech Revolution*. The World Bank.

Demirgüç-Kunt, A, L Klapper, D Singer, S Ansar, and J Hess. 2017. “Measuring Financial Inclusion and the Fintech Revolution.” *The Global Findex Database, World Bank Group*.

Edwards, Jeffrey R, and Richard P Bagozzi. 2000. “On the Nature and Direction of Relationships Between Constructs and Measures.” *Psychological Methods* 5 (2). American Psychological Association: 155.

Géron, Aurélien. 2019. *Hands-on Machine Learning with Scikit-Learn, Keras, and Tensorflow: Concepts, Tools, and Techniques to Build Intelligent Systems*. O’Reilly Media.

Glorot, Xavier, and Yoshua Bengio. 2010. “Understanding the Difficulty of Training Deep Feedforward Neural Networks.” In *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, 249–56.

Gosset, WS. 1908. “The Probable Error of a Mean” *Biometrika*, Vol. 6, N. 1, Pp. 1-25, Originally Published Under the Pseudonym.” Student.

Griffoli, Mr Tommaso Mancini, Mr Maria Soledad Martinez Peria, Mr Itai Agur, Mr Anil Ari, Mr John Kiff, Ms Adina Popescu, and Ms Celine Rochon. 2018. *Casting Light on Central Bank Digital Currencies*. International Monetary Fund.

He, Mr Dong, Mr Ross B Leckow, Mr Vikram Haksar, Mr Tommaso Mancini Griffoli, Nigel Jenkinson, Ms Mikari Kashima, Tanai Khiaonarong, Ms Celine Rochon, and Hervé Tourpe. 2017. *Fintech and Financial Services: Initial Considerations*. International Monetary Fund.

IMF. 2009. “International Transactions in Remittances: Guide for Compilers and Users.” IMF Washington, DC.

Ji, Suhwan, Jongmin Kim, and Hyeonseung Im. 2019. “A Comparative Study of Bitcoin Price Prediction Using Deep Learning.” *Mathematics* 7 (10). Multidisciplinary Digital Publishing

Institute: 898.

Kingma, Diederik P, and Jimmy Ba. 2014. "Adam: A Method for Stochastic Optimization." *arXiv Preprint arXiv:1412.6980*.

MacKinnon, James G. 1991. "Critical Values for Cointegration Tests." In *Eds.), Long-Run Economic Relationship: Readings in Cointegration*. Citeseer.

Montgomery, DC, EA Peck, and GG Vining. 2001. "Introduction to Linear Regression Analysis, Htujohn Wiley & Sons." *Inc. UTH, New York, NY*, 292–96.

Neel Mehta, Parth Detroja, Adi Agashe. 2019. "Bubble or Revolution. the Present and Future of Blockchain and Cryptocurrencies." In, 67–69. Paravane Ventures.

Ng, Serena, and Pierre Perron. 1995. "Unit Root Tests in Arma Models with Data-Dependent Methods for the Selection of the Truncation Lag." *Journal of the American Statistical Association* 90 (429). Taylor & Francis: 268–81.

Pascanu, Razvan, Tomas Mikolov, and Yoshua Bengio. 2013. "On the Difficulty of Training Recurrent Neural Networks." In *International Conference on Machine Learning*, 1310–8.

Sánchez de Diego Martínez-Cabrera, Jaime. 2014. "Bitcoins¿ Revolución O Historia?"

Shmueli, Galit, and others. 2010. "To Explain or to Predict?" *Statistical Science* 25 (3). Institute of Mathematical Statistics: 289–310.

Struga, Kejsi, and Olti Qirici. 2018. "Bitcoin Price Prediction with Neural Networks." In *RTA-Csit*, 41–49.

Valencia, Franco, Alfonso Gómez-Espinosa, and Benjamín Valdés-Aguirre. 2019. "Price Movement Prediction of Cryptocurrencies Using Sentiment Analysis and Machine Learning." *Entropy* 21 (6). Multidisciplinary Digital Publishing Institute: 589.

Wu, Shaohua, TJ Harris, and KB McAuley. 2007. "The Use of Simplified or Misspecified Models: Linear Case." *The Canadian Journal of Chemical Engineering* 85 (4). Wiley Online Library: 386–98.

Zhu, Dongming, and Victoria Zinde-Walsh. 2009. "Properties and Estimation of Asymmetric Exponential Power Distribution." *Journal of Econometrics* 148 (1). Elsevier: 86–99.

Zimmerman, Peter. 2020. "Blockchain Structure and Cryptocurrency Prices." Bank of England Working Paper.